Enshrouded Exoplanetary Systems

Thesis

The Open University

Enshrouded Exoplanetary Systems

Submitted for the degree of Doctor of Philosophy

in Astronomy

Daniel Staab

The

Open University

May 2017

Abstract

This thesis is based on the hypothesis that stars hosting short period planets can show anomalously low apparent stellar activity values caused by planetary mass-loss enshrouding the system. The resulting circumstellar absorption depresses the cores of the Ca II H & K lines, where the activity index $\log(R'_{HK})$ is measured. I searched for this effect in a sample of ~2700 stars, identifying 39 objects with $\log(R'_{HK})$ below the main sequence basal limit, and 6 Hyades and Pleiades stars with activity levels well below the cluster distributions. I describe a radial velocity programme to search for short period planets orbiting these stars. Because nearby bright stars are targeted, their mass-losing companions will offer key opportunities to probe planetary composition through transmission spectroscopy. I present four planet candidates discovered through this programme orbiting HD11231 and HD38677. The former is a γ-Doradus pulsator hosting a Saturn-mass planet in a 6 day orbit, and the latter is a compact system containing planets with ~18, 13 and 3 Earth masses and periods between 20 and 1.7 days.

In addition, the Robert Stobie Spectrograph was calibrated to measure $\log(R'_{HK})$ values. The activity index of the known Hot Jupiter hosts WASP-43, WASP-51/HAT-P-30, WASP-72 & WASP-103 was measured to search for anomalous values caused by the companions. The activity level of WASP-43 is extremely high in absolute terms and relative to its age and X-ray emission. The activity of WASP-103 is marginally higher than expected from the system age. In both cases, star-planet interactions may enhance Ca II H & K emission. WASP-51/HAT-P-30 and WASP-72 show anomalously low activity; the latter falls below the basal envelopes of main sequence and evolved stars. I highlight that a quarter of transiting, short period planets hosts with published activity data exhibit anomalously low activity. These 22 systems include Hot Jupiters and low mass companions.

Acknowledgements

I am very grateful to my family, especially Mum and Dad: without your emotional and practical support this would never have been possible, thanks for everything you have done over the years! Thanks go to my supervisor Carole Haswell, who provided me with this great opportunity of doing exoplanet science and innumerable hours of excellent supervisor support, advice and help. I greatly appreciate the contributions of many colleagues throughout this PhD, especially John Barnes, Guillem Anglada-Escude, Zaira Berdinas, Gareth Smith, Luca Fossati, and Geoff Bradshaw.

Thanks to the STFC for funding my work, and ESO & OPTICON for providing plentiful telescope time, excellent logistics and observatory staff. It has been great spending this PhD era with my lovely housemates Loua, Patrizia, Kurijn, Kuba, Tom & Sam; a string of great office mates; and the many wonderful people I've met in and around the Open University. There are far too many of you to list here, but thanks to all of you I will look back at many fond memories of my time here!

I would like to include a special mention of the great people at the University of Leicester, and my secondary school teacher Rita Sommer, who were all important parts of my journey through science and astronomy.

I dedicate this thesis to the memories of Franziska Staab, Maddy Orford, Pi Tuk, Jakubko Sedivy and Ann Spiller who passed away far too early.

Contents

List of Figures

List of Tables

Chapter 1

An Introduction to Exoplanets, Planetary Mass Loss and Stellar Activity

”Exploration is in our nature. We began as wanderers, and we are wanderers still. We have lingered long enough on the shores of the cosmic ocean. We are ready at last to set sail for the stars.“

-Carl Sagan, 1980

1.1 Exoplanets: All these worlds are yours ...

Starting with the very first discoveries of exoplanetary systems at the end of the 20th century, observations have revealed populations of planets that are surprisingly diverse and arguably even extreme when compared to the solar system. The somewhat geocentric, or rather solar-centric viewpoint underlying this surprise is understandable. Our Solar System seemed to be a sensible, orderly arrangement of things: the inner rocky planets, followed by the gas and ice giants beyond the snow line, all on respectably low-eccentricity orbits around a stable, middle-aged star. There have now been two decades of exoplanet discovery, of astronomy's first mapping of Worlds in the Galactic neighbourhood. A picture has emerged of planetary systems existing in almost any configuration that is both allowed by the basic laws of physics and compatible with our astronomical detection capabilities. Planets have been discovered orbiting pulsars (Wolszczan & Frail 1992); multiple stars (Doyle et al. 2011) and objects ranging from giant, 3 solar mass stars (Sato et al. 2012) to ultra-cool dwarfs with a trifling 0.08 solar masses (Gillon et al. 2017). They have been found on highly eccentric, comet-like orbits (Jones et al. 2006), and in tightly packed systems where half a dozen planets are huddled within the orbital distance of Mercury (Lissauer et al. 2011). Gas giants were detected orbiting only two stellar radii above the photosphere of their parent stars (Hebb et al. 2009). And the current catalog of thousands of known exoplanets highlights that the most frequently discovered "super-Earths" or "mini-Neptunes" (1-4 times the size of Earth, Batalha 2014), are absent from our Solar System.

Due to the difficulty of detecting small, low-mass planets with periods longer than a few days, it is still largely unknown how common or uncommon systems like the Solar System really are. For now, this question - and the detection of just a single direct Earth analogue - remain at the edge of observational astronomy's

reach. Roadmaps towards an answer are being rigorously planned (e.g. Mayor et al. 2014). Discoveries of potentially habitable, yet distinctly non-Earth-like planets such as those transiting a nearby ultra-cool dwarf star (Gillon et al. 2017) are already providing tantalising opportunities for follow-up studies. And for the first time plans are being drawn to send interstellar probes to such planets - not in the distant future, but within a generation (Hoang et al. 2017), emboldened by the discovery of our nearest neighbour Proxima b orbiting only 1 parsec away (Anglada-Escudé et al. 2016a).

It is worth contemplating just how diverse the zoo of known planets has become, despite the limited information we have access to. The diversity described above encompasses the planetary mass and size and the stellar/orbital parameters, which determine the irradiation level. Little else is known about the vast majority of exoplanets. Characterisation of the composition and atmospheres of these planets is still at an early stage (Burrows 2014), while information on their surfaces is beyond the reach of current telescopes and instrumentation. I will return to the topic of characterisation in Section 1.1.2 after describing the exoplanet discovery methods used (Section 1.1.1). I describe a particularly extreme and interesting category of planets in Section 1.2. Section 1.3 describes the basics of stellar activity, and its possible connections with exoplanets. Finally, the main research questions of my thesis are described in Section 1.5.

1.1.1 Exoplanet detection methods

The use of transits and radial velocities (RV) are by far the most productive exoplanet discovery methods. I describe these in more detail below as they are the techniques relevant to this thesis. For completeness, I will briefly list other techniques in use. I refer the reader to Wright & Gaudi (2012) for an in-depth general discussion of exoplanet detection methods. My discussion here follows their work in part, using other references as indicated.

Direct Imaging:

The conceptually most obvious method is extremely challenging in practise: even in the best cases, flux contrast ratios between stars and planets of 10^{-4} - 10^{-6} at separations of less than 1'' must be overcome. A modest number of gas giants on wide orbits have been discovered this way to date (e.g. Lannier et al. 2016). The technique is currently able to detect such planets in young systems, when higher thermal emission from the planets provides a better contrast ratio than for older systems. In the near future, a crucial exception may be provided by a *Breakthrough Initiatives* project[1], which aims to be sensitive to Earth-like planets in the habitable zones of the nearest sun-like stars, α-Centauri A and B. The key advantage of direct imaging is that detailed spectroscopic characterisation of the light emitted and reflected by the planets becomes possible.

Microlensing:

Microlensing is a relativistic effect of the bending of light rays by a foreground mass, where its relative motion to a background source causes an increase in brightness. This requires a rare chance alignment of foreground lens and the background object. The vast majority of observations are therefore large photometric surveys, targeting the dense Galactic bulge (e.g. Gaudi 2012).

[1] http://breakthroughinitiatives.org/News/8

Around 50 planets, ranging from Earth-mass to super-Jupiters have been discovered through microlensing to date[2]. An intriguing aspect of the microlensing method is that it can discover free-floating exoplanets (without a host star), as it relies only on the gravity of the lensing object. Preliminary evidence (Sumi et al. 2011) is consistent with a Galactic population of free-floating planets outnumbering main sequence stars two to one. This controversial finding is now being tested by space-based observations (Penny et al. 2017).

Pulsar timing:

This method is an oddity in the history of exoplanet science. It provided the first confirmed exoplanets (Wolszczan & Frail 1992), which were also very low-mass (~3 Earth masses), but it has led to the discovery of only ~20 planets since then[2]. It is unclear how these exotic planets either survive the super nova stage creating the pulsar, or how they form subsequently. The pulsar technique relies on precision arrival time measurements of the pulses. Periodic variations of the intrinsically stable pulsar "clock" are interpreted as motion of the pulsar around a barycentre[3] in response to orbiting companions.

Astrometry:

With highly precise stellar parallax measurements, it is possible to directly measure the on-sky motion of stars around the barycentre of a planetary system. The method is most sensitive to massive planets in wide orbits around nearby stars. Only a single exoplanetary system is currently considered to be a valid astrometric detection[2]. However, the ongoing extremely precise parallax measurements by the GAIA space telescope are expected to reveal ~20,000 long period gas giants (Perryman et al. 2014).

[2]http://exoplanets.eu; accessed 2017-04-13
[3]the system's centre of mass

Transit

To date, the vast majority (~2700, equivalent to 75%) of exoplanets have been discovered via transit[2]. A Jupiter-sized planet orbiting a sun-like star causes a periodic 1% dip in stellar flux, which can be readily detected with small, ground-based telescopes. Clearly, the random orbital inclination (i) relative to the observer's line of sight has to be aligned correctly for transits to occur. The probability for such an alignment drops off as a^{-1}, where a is the planet's semi-major axis. Even for the favorable case of close-in (a ~0.05 AU) Hot Jupiter (HJ) planets, the probability is only of order 10%, and such objects exist around only ~1% of stars (Wang et al. 2015). Note that there is no unique definition of the HJ category. Throughout this thesis, I adopt the rough criterion of planet mass $M_p > 0.5$ M_J and $P < 10$ d, as in the Exoplanet Orbit Database (Han et al. 2014). Given the low transit probabilities for individual systems, large surveys are carried out, monitoring 10^3-10^7 stars. With accurate characterisation of the host star and its radius (R_*), the planetary radius (R_p) is obtained directly from the transit depth.

Ground-based efforts like the Wide Angle Survey for Planets (WASP, Pollacco et al. 2006) have revealed hundreds of planets with approximately Jovian radii[2]. To detect the much smaller transit depths (0.01%) of Earth-sized planets around sun-like stars, space-based observatories above Earth's atmospheric noise are needed. The Kepler space telescope has revolutionized exoplanet science by discovering ~2300 confirmed transiting planets to date[2], and establishing population statistics across a wide region of parameter space (e.g. Batalha 2014).

A complication in transit surveys is the array of other astrophysical signals that can mimic exoplanetary transits, i.e. false positives. Common false positives are unfortunate alignments with eclipsing binary stars and transits by planet-sized stars. To identify such cases, a barrage of tests can be carried out

on the stellar light curves, in addition to high contrast imaging of the candidate planet host star and its vicinity. Often the final confirmation of transit candidates is done via radial velocity (RV) measurements (see below). The development of statistical tools has meant that such RV follow-up is not always required for confirmation (e.g. Lissauer et al. 2014).

Key points that I will return to in this chapter are that a) the planet's inclination i can be measured from the transit light curve, and b) transmission spectroscopy via transits are a vital probe for exoplanet characterisation.

Radial Velocity (RV) method

This method exploits the motion of a planet host star around the system's barycentre in response to one or more orbiting companions, as done in astrometry. Rather than measuring on-sky position changes of the star, the method searches for periodic variations in the line-of-sight velocity of a star, i.e. its radial velocity (RV). This is done via the Doppler effect. In essence, a series of stellar spectra are recorded over time, and the observed wavelength shift at time t, relative to a reference wavelength λ_0, is converted to the corresponding RV:

$$\mathrm{RV} = c\,\frac{\lambda(t) - \lambda_0}{\lambda_0}, \tag{1.1}$$

where c is the speed of light. Implementing this in practice for low-amplitude planetary signals is an extremely challenging instrumentation and data analysis problem. A detailed discussion of these aspects will be provided in Chapter 4. Note that important relativistic terms are ignored in Equation 1.1, as they are constant in time for the precision and timescales of interest and therefore irrelevant (e.g. Lovis et al. 2010).

The radial velocity of a star with mass M_* orbiting the barycentre of a system with a single planet of mass M_p at time t can be described by (Wright & Howard 2009):

$$\mathrm{RV}(t) = \left(\frac{2\pi G}{P}\right)^{1/3} \left(1-e^2\right)^{-1/2} \frac{M_\mathrm{p} \sin i}{(M_* + M_\mathrm{p})^{2/3}} \Big[\cos(f(t)+\omega) + e\cos\omega\Big] + \gamma. \quad (1.2)$$

Equation 1.2 includes the gravitational constant G, the orbital period P, the orbital eccentricity e and the argument of periastron ω. The latter specifies the angle on the orbital plane with respect to the plane of the sky where the star is at the pericentre[4] of its orbit. Note the planet's argument of periastron is at ω+180 degrees. The true anomaly $f(t)$ is the angle between the pericentre and the position of the star at time t, as seen from the focus of its orbital ellipse. It can be defined implicitly via the eccentricity anomaly $E(t)$ and mean anomaly $M(t)$:

$$\tan\left(\frac{f(t)}{2}\right) = \sqrt{\frac{1+e}{1-e}}\, \tan\left(\frac{E(t)}{2}\right) \quad (1.3)$$

$$M(t) = E(t) - e\sin E(t) = \frac{2\pi(t - t_\mathrm{peri})}{P}, \quad (1.4)$$

where t_peri is the time of pericentre passage. In the literature, orbital elements are often reported by specifying the value of the mean anomaly at the reference epoch t_0, designated M_0 (e.g. Tuomi et al. 2013, Anglada-Escudé et al. 2013). γ is a constant reference velocity (approximately that of the system's barycentre), which can be subtracted off as the mean velocity of the RV timeseries. It is worth emphasising that only **relative** RVs (to γ) are required and of interest for planet detection. In most cases the measured γ value is not equal to the precise **absolute**, astrophysical, mean velocity of a star. The latter is extremely difficult to determine at high precision due to both astrophysical and instrumental systematic effects (e.g. Lovis et al. 2010). These effects are constant in time, so

[4] the point of closest approach to the focus of the orbital ellipse

they do not impede the measurement of high-precision relative RVs. However, datasets of a given star obtained with different instruments typically show a constant, non-negligible offset between their γ values. An instrumental upgrade can also cause a one-off systematic jump in γ (Lo Curto et al. 2015). When datasets are combined for planet detection, these issues are factored into the analysis.

Equation 1.2 represents a single Keplerian signal, in a multi-planet system the superposition of many such signals governs the stellar RV timeseries[5]. In the exoplanet literature, it is convention to quote the semi-amplitude K of the RV curve, as can be seen from Equation 1.2, this is given by

$$K = \left(\frac{2\pi G}{P}\right)^{1/3} \left(1 - e^2\right)^{-1/2} \frac{M_\mathrm{p} \sin i}{(M_* + M_\mathrm{p})^{2/3}}. \tag{1.5}$$

For circular orbits $e = 0$ and ω and t_peri are formally undefined. In such cases various conventions are used in the literature. In this thesis I will replace t_peri in Equation 1.4 by t_0, the first epoch of the relevant timeseries. $M(t)$, $f(t)$ and $E(t)$ are then equal, and still defined via Equations 1.3 and 1.4 as $2\pi(t - t_0)/P$. The stellar RV curve is then fully specified via t_0 and the best-fit values for ω, K, P and γ (c.f. Equations 1.2 and 1.5). ω and t_0 determine the phasing of the RV curve in the circular case.

From an observed RV timeseries the orbital elements of the star and planets can be deduced via Equation 1.2. M_* is obtained indirectly through analysis of a combination of the stellar spectrum, photometry, parallax and stellar evolution models, or through astroseismology. Unless the planet transits, the inclination i remains unknown and only $M_\mathrm{p} \sin i$ is obtained. It is convention to report the minimum mass value, labelled as $M_\mathrm{p} \sin i$, which is equal to the true mass if $i = 90$ degrees[6]. This $\sin i$ ambiguity is a challenge for the characterisation of individual planets, but is not a significant problem for statistical studies of

[5] in many cases, complications arise from gravitational perturbation of multi-planet orbits
[6] an exactly edge-on orbit as viewed from the observer

planet populations. It can be shown that for randomly distributed i values, the probability distribution of $\sin i$ is biased towards edge-on orbits. The average value is equal to $\pi/4 = 0.79$, and the probability that $\sin i > 0.5$ is 87% (e.g. Lovis et al. 2010). For the reference case of i = 90 degrees and a M_* = 1 $M_\odot$ star, a direct 1 Jupiter mass analogue (with P = 12 years) induces a semi-amplitude of 13 $m\,s^{-1}$. Hot Jupiters can induce K ~100 - 1000 $m\,s^{-1}$, and an Earth mass planet in a 1 year orbit induces K ~0.1 $m\,s^{-1}$. Earth mass planets in the habitable zones of low-mass M-dwarf stars have short periods due to their hosts low luminosity, and can induce K ~1 $m\,s^{-1}$ (e.g. Anglada-Escudé et al. 2016a).

Discovery via the RV method dates back to the first exoplanet found around a sun-like star (51 Peg) by Mayor & Queloz (1995). Announcements of new worlds were dominated by RV results throughout the nineties and noughties, until the transit method came into its own[7], in particular due to the Kepler mission. The initial exoplanet haul consisted predominantly of short-period and eccentric gas giants. With improving instrumentation, RV precision, and longer survey baselines, the accessible parameter space for RV discoveries widened considerably. There are now frequent detections of signals with periods ranging from less than 1 day to 15 years and semi-amplitudes as low as 1-2 $m\,s^{-1}$. As shown by Figure 1.1, this corresponds to the discovery of planets down to 1 $M_\oplus$ in short period orbits, while the majority of RV exoplanets fall in the super-Earth and gas giant categories. Planets detected via both RV and transit signals are particularly valuable. The $\sin i$ ambiguity is resolved, and the measured mass and radius allow the bulk density to be calculated. This gives important constraints on the structure of the planets, e.g. whether a significant volatile envelope is present or not. In the absence of detailed compositional information, such exoplanets are at least characterised to first order.

[7] http://exoplanetarchive.ipac.caltech.edu/exoplanetplots/

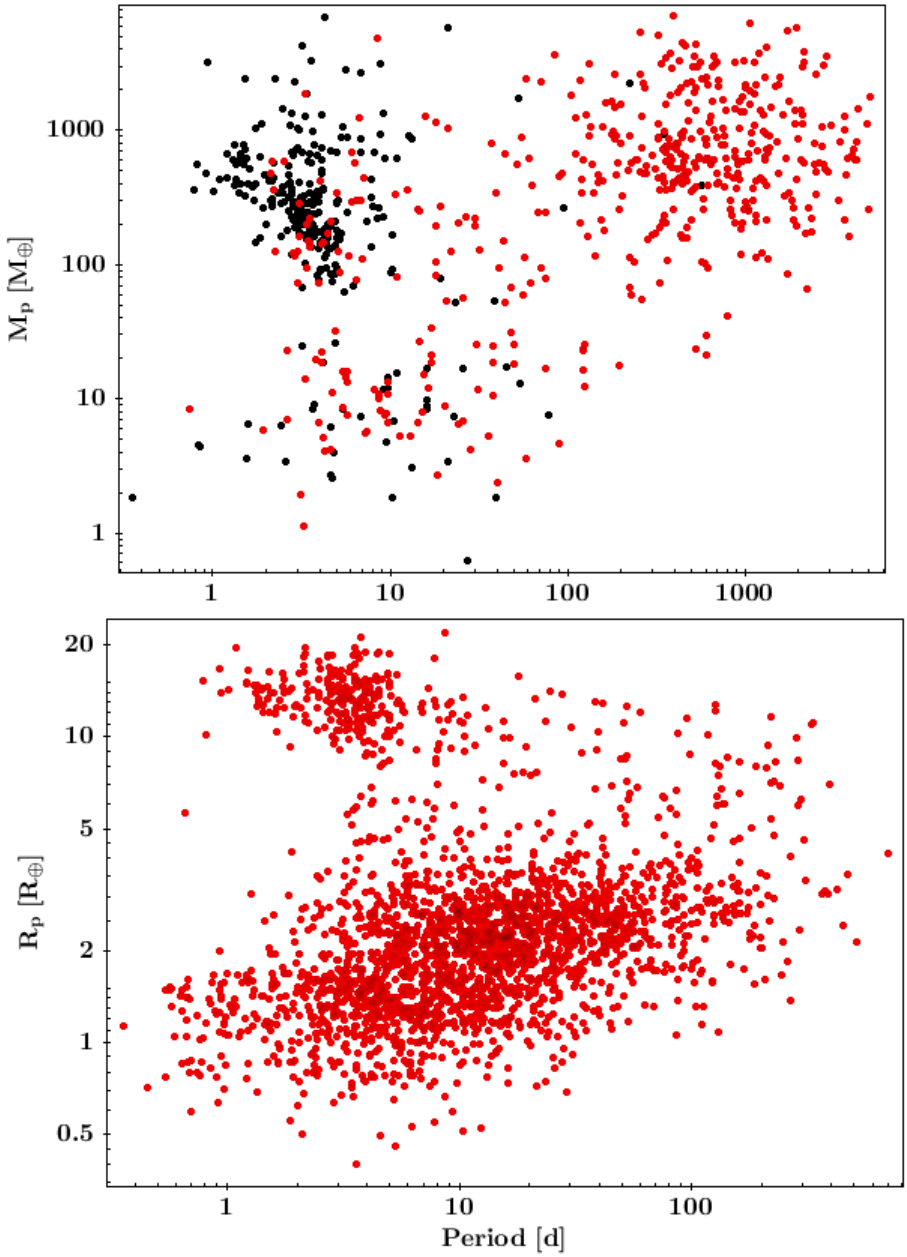

Figure 1.1: **Top:** The mass-period distribution of the known exoplanet population, for planets with a mass constraint from RV measurements. For planets discovered by the transit method (black) the true mass is known, while for planets with only RV data (red) the minimum mass $M_\mathrm{p} \sin i$ is plotted. Obvious clustering occurs for long-period gas giants, Hot Jupiters with P < 10 days, and short to medium period super-Earths & Neptunes (1 - 30 $M_\oplus$). The clustering is in part due to selection bias, but the so-called "sub-Jovian desert" at short periods ($\lesssim$ 3 days) is a real feature of the exoplanet population (Szabó & Kiss 2011). **Bottom:** The radius-period distribution of the known transiting exoplanets. The "sub-Jovian desert" is also clearly visible in this sample (Beaugé & Nesvorný 2013).

As is the case for transits, RV measurements are affected by false-positive signals. This will be discussed in detail in Section 4.2. The RV method is used extensively in this thesis. I provide a detailed discussion of the instrumentation, data reduction and analysis techniques involved in Chapter 4.

1.1.2 Measuring exoplanets' atmospheric composition

The existence of Hot Jupiters provided astronomers with a fortuitous early opportunity to attempt the characterisation of distant planets, orbiting in very different environments than the planets of our own system. Key characterisation work has been done via transmission spectroscopy, i.e. differencing the stellar spectrum during and away from transit to reveal absorption features in the planetary atmosphere. HJs are a best-case scenario for this work; with frequent transits and an atmospheric cross-section that is large enough in several wavelength regions to be detectable with current facilities. The dense, lower atmospheres of exoplanets can be probed in strong optical absorption lines and in infrared bands of abundant molecules. The first atmospheric detection was achieved via the strong sodium (Na) D doublet (Charbonneau et al. 2002) with space-based observations. Redfield & Linsky (2008) provided the first ground-based detection of Na, and Cauley et al. (2015) detected an absorption signature in the Hydrogen Balmer lines. In the infrared, absorption spectra consistent with water vapour are the most common (see Iyer et al. 2016 for a recent overview). Even for HJs, measurements of the lower atmosphere are challenging: the absorbing atmospheric annulus typically causes signals at the 10^{-3} - 10^{-4} level. In many cases, "flat" spectra are obtained instead, and interpreted as clouds or haze muting any gas-phase absorption lines (e.g. Wakeford et al. 2017).

The upper atmospheres of some short period planets provide much more favorable opportunities for compositional measurements. Irradiating flux from the extremely close-by host stars can have dramatic effects on planets like Hot Jupiters. The intense, high-energy X-ray and ultraviolet (UV) stellar radiation field can drive high rates of planetary mass loss. Such outflows of diffuse gas exceeding escape velocity were first observed for the Hot Jupiter HD 209458b in the seminal work of Vidal-Madjar et al. (2003). A dramatic 15% Ly-α transit depth was seen, while the optical transit depth is only 1.5%. The Ly-α transit must be caused by an escaping Hydrogen cloud with an effective radius of 4.3 Jupiter radii (R_J), well above the planet's radius of 1.4 R_J and its Roche Lobe (3.6 R_J). Further observations in the far UV were inspired by the discovery, for instance, of HD189733 (Lecavelier des Etangs et al. 2010, Lecavelier des Etangs et al. 2012) and 55 Cnc (Ehrenreich et al. 2012). A recent, spectacular example is the 56% Ly-α transit depth detected for the warm Neptune GJ 436b (Ehrenreich et al. 2015). The extended, asymmetric and deep transit implies that a very substantial comet-like hydrogen tail surrounds the planet and extends out to several stellar radii (Figure 1.2).

Large-scale hydrodynamic escape of the upper atmosphere of irradiated short-period gas planets has been modelled extensively (e.g. Lammer et al. 2003; Bisikalo et al. 2013; Matsakos et al. 2015). Heavy atomic species are entrained in these prodigious outflows of hydrogen (Linsky et al. 2010; Fossati et al. 2010a; Vidal-Madjar et al. 2013; Ben-Jaffel & Ballester 2013). Systems with such significant mass loss are key targets for transmission spectroscopy. From the point of view of an observer, the escaping gas covers a much larger area of the star than the thin annulus of a typical exoplanetary atmosphere and produces a much larger absorption signal. The physical parameters of the low density outflows mean UV observations are the ideal probe for characterisation (Fossati et al. 2015a). A large number of chemical species have strong spectral lines in this wavelength region and can be detected (Haswell et al. 2012).

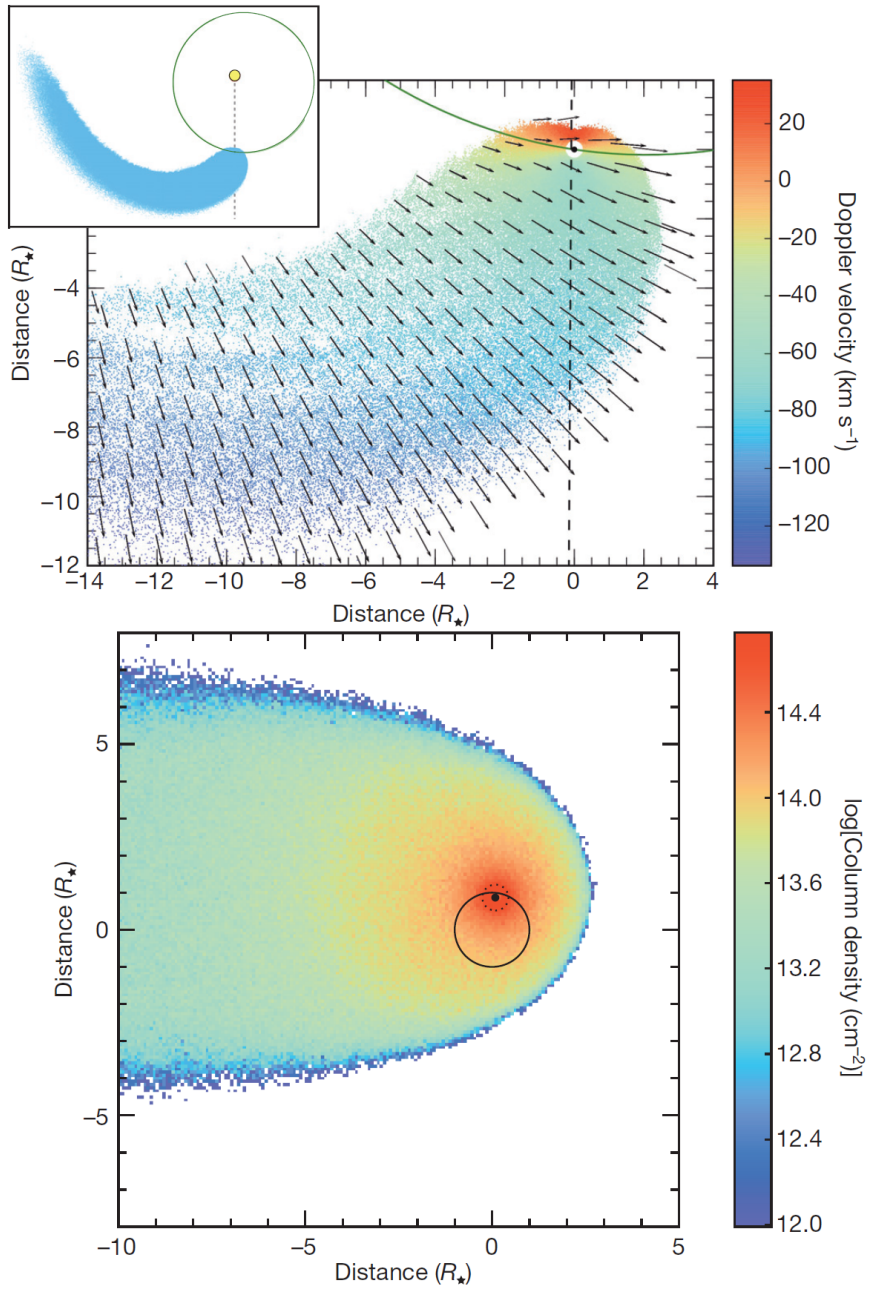

Figure 1.2: Visualisations of the enormous comet-like Hydrogen tail created by mass loss from the Neptune-mass planet GJ 436b. Plots are taken from Ehrenreich et al. (2015), and are outputs from a three-dimensional model consistent with the Ly-α transit observed in that work. **Top panels:** A face-on view of the system's orbital plane, showing the orbital path of the planet (green) around GJ 436 (yellow; to scale) and the line-of-sight of the transit observations (dashed). The inset shows the entire cometary tail, while the detailed view highlights the cloud's velocity vectors in the frame of the star. The projected velocities along the line-of-sight are colour-coded and match the observed velocities of Ly-α absorption. **Bottom panel:** Edge-on view of the system during the transit of the planet (black dot) across the stellar disk (black circle). The Hydrogen column density of the escaping material is colour-coded. Dense regions of this cloud extend far beyond the planet's Roche Lobe (dashed) and the stellar disk.

1.2 Ultra-Short Period Planets

Independently of the efforts described in the previous section, strong evidence for large-scale mass-loss was discovered in a population of so-called ultra-short period planets (USP, defined as $P < 1$ day). This term was coined by Sanchis-Ojeda et al. (2014), who conducted a thorough, dedicated transit search for such objects via Kepler data and inceased the number of USP candidates to ~100. Orbital periods as short as 0.17 days (4 hours) were observed. This is close to the shortest physically allowable periods. From Kepler's law, the orbital period around a K1 star (M_*= 0.8 $M_\odot$) which corresponds to a semi-major axis equal to the stellar radius (0.85 R_*) is 2.4 hours. Clearly, a planet cannot orbit within the stellar radius. Additional restrictions are imposed by the Roche limit of tidal disruption. Rappaport et al. (2013) showed that for the most extreme USPs, this constrains the planetary density to values consistent with an iron-dominated composition.

Common properties observed or inferred for this distinct category of planets are (Sanchis-Ojeda et al. 2014):

- They are rare objects, found around ~0.5% of stars.
- Frequent occurrance in multi-planet systems (in contrast to HJs), with additional transiting companions found at periods of 1-50 days.
- Planetary radii are almost exclusively < 2 $R_\oplus$, in contrast to longer period transiting planets.
- Very high irradiation levels (10^2- 10^4 times the flux received by Earth), that could have removed any original volatile envelope for a wide range of plausible model constraints. This is consistent with the small radii observed.

- A likely rock-dominated composition for the majority of USPs. Surface temperatures of $\gtrsim$ 2000 K would vaporise any silicate minerals. A wide range of compositions for exotic silicate-vapor atmospheres, and surface lava oceans could occur (e.g. Miguel et al. 2011, Kite et al. 2016).

The formation and orbital evolution of USPs are still poorly understood (e.g. Lopez & Rice 2016, Adams et al. 2016 and references therein). Due to their very short periods, they are expected to have undergone tidal locking and orbital circularisation.

A spectacular discovery of the catastrophically disintegrating, rocky planet Kepler-1520b (originally designated KIC 12557548b) by Rappaport et al. (2012) raised the prospect of studying the composition of material sublimated from the surface of USPs. The photometric transit signals detected were consistent with a time-variable dust cloud around the planet, associated with direct sublimation of the surface (Budaj 2013; van Lieshout et al. 2014; Bochinski et al. 2015). Theoretical models indicate that this phenomenon is a final, short-lived phase of large-scale mass loss undergone by the most extreme, low-mass USPs (Perez-Becker & Chiang 2013). Depending on the system's parameters, the complete disintegration of such a body during the main sequence life-time of its host star is possible. Two additional disintigrating rocky planets have since been discovered in Kepler data: KOI-2700b (Rappaport et al. 2014) and K2-22b (Sanchis-Ojeda et al. 2015). The vast majority of USPs show no sign of dusty outflows in their transit signals. Investigating spectroscopic signals of gas-dominated mass loss is impractical for the currently known population due to the host stars' faintness[8]. This issue limits the scope of follow-up studies of all known catastrophically disintegrating bodies, and of most USPs in general. Finding and characterising rocky USPs with significant mass loss rates around bright stars could therefore lead to a step change in comparative planetology.

[8]This is due to the Kepler sample (from which most of them originate) being biased towards large stellar distances.

1.3 Stellar activity

1.3.1 History and definitions of activity measurements

Observations of stellar activity and the characteristic H & K lines of ionized Calcium (hereafter Ca II H & K) are a cornerstone of this thesis. Before I discuss their relevance to exoplanetary systems, I will give a brief account of their study and "classical" relevance in stellar astrophysics. A thorough review covering the extensive literature on the nature and observation of stellar activity can be found in Hall (2005). In brief, the stellar magnetic surface flux drives nonthermal heating of the chromosphere, leading to emission in several lines, including the optical Ca II H & K doublet and the ultraviolet Mg II h & k lines. Stars with convective envelopes exhibit a wide range of magnetic activity levels, and consequently vary substantially in their chromospheric emission. To first order, the stellar activity level depends on the stellar rotation rate and therefore age, with a particularly rapid decrease during spin-down over the first ~ 1 Gyr (e.g. Mamajek & Hillenbrand 2008 and Pace 2013). Measurements of chromospheric emission are thus used as a stellar age indicator.

The strong Ca II H & K resonance lines cause very deep photospheric absorption features in late-type stars. Chromospheric emission reversals are located at the lines' cores, super-imposed on the photospheric flux (c.f. Figures 1.3 and 1.4). Measuring this emission is historically very well established as a stellar (chromospheric) activity indicator and was pioneered by a purpose-built photometer at the Mount Wilson observatory in the 1960s (Wilson 1968, Duncan et al. 1991). As active regions rotate with the stellar surface, emission levels are modulated by a star's rotation period. Over long baselines, the emission traces stellar activity cycles, which was the focus of the multi-decade Mount Wilson program. The bandpasses of the photometer used by this survey are shown in Figure 1.3, and gave rise to the S-index as the standard Ca II H & K activ-

ity proxy to this day. Details of the bandpass boundaries and full-width half maxima (FWHM) are reported in Section 3.3 of this thesis. The dimensionless S-index is simply the ratio of integrated fluxes in the core bandpasses, divided by the integrated flux in the continuum bandpasses. This ratio is often multiplied by a historical scaling factor. However, this becomes obsolete if the mean bandpass fluxes are used throughout instead of integrated fluxes (Lovis et al. 2011). I will adopt this approach in my thesis, and I therefore omit discussing the origin and details of the scaling factor.

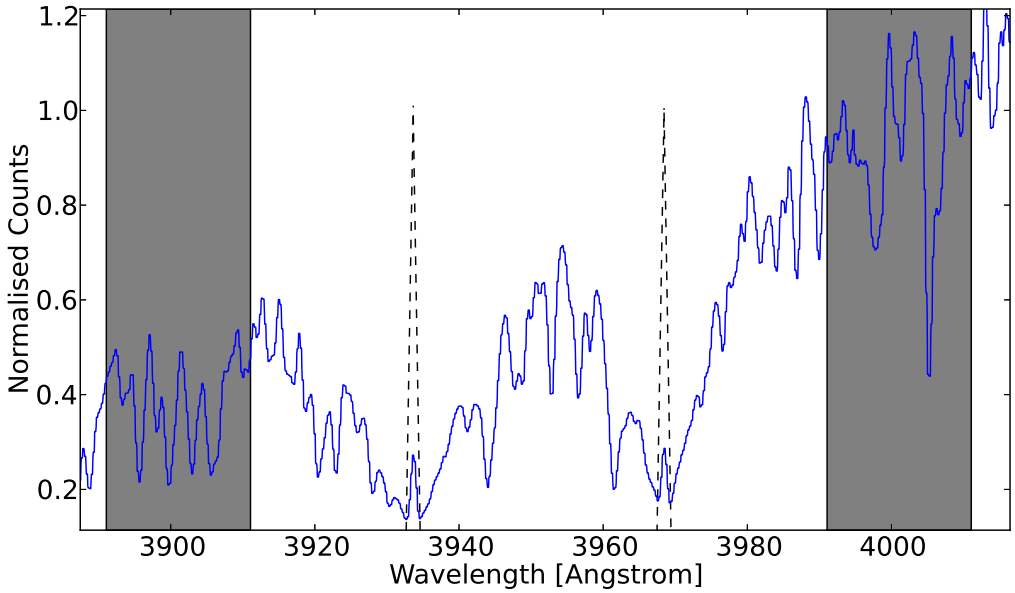

Figure 1.3: The Ca II H & K lines of a star with significant chromospheric emission, as seen in a low resolution spectrum. The continuum windows (greyed) and triangular core bandpasses used to calculate the S-index are highlighted. This figure is repeated in Chapter 2 of this thesis, where I report the details of the observation.

Following the Mount Wilson survey, S-values have been measured for thousands of stars (e.g. Isaacson & Fischer 2010; Jenkins et al. 2011; for a detailed account see Section 2.1). In these "second generation" studies, S-values are extracted from medium to high resolution spectra via synthetic photometry, and then calibrated onto the Mount Wilson scale by the observation of a set of common calibrator stars known to have particularly stable activity levels.

This calibration removes systematic effects associated with the different spectrographs used, allowing stars to be compared on a consistent scale across large samples and long baselines.

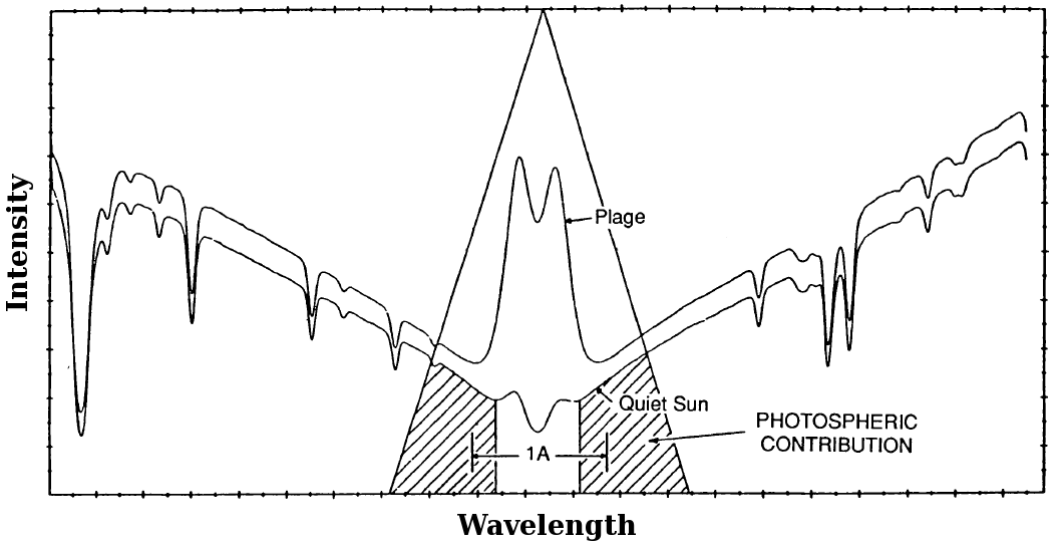

Figure 1.4: Illustration of the Ca II K line core profile at high resolution, adapted from Duncan et al. (1991). The triangular core bandpass includes chromospheric and photospheric contributions. The two spectra are offset for better visibility, and are both solar observations. They show the difference between the average sun state and the spectrum of a plage (active region). This is analogous to a comparison between inactive and highly active stars with otherwise identical spectral properties.

The S-value is sufficient to trace the activity of individual stars. However, its dependence on stellar colour makes a different metric necessary when comparing stars with different spectral types. Since the work of Noyes et al. (1984), this has predominantly been the dimensionless $\log(R'_{\mathrm{HK}})$ index. It accounts for the colour-dependence of the continuum bandpasses' flux and subtracts out the (colour-dependent) photospheric contribution to the core bandpasses, illustrated in Figure 1.4. This is done via empirically established relations that use the widely-measured stellar $B-V$ colour-index.

The relation defining $\log(R'_{\rm HK})$, following Noyes et al. (1984), is

$$R'_{\rm HK} = R_{\rm HK} - R_{\rm phot}. \tag{1.6}$$

All quantities in Equation 1.6 are normalised by the total bolometric flux of the star and are dimensionless. $R_{\rm HK}$ is the summed normalised flux in the H & K core bandpasses. $R'_{\rm HK}$ and $R_{\rm phot}$ are the normalised chromospheric and photospheric contributions to $R_{\rm HK}$ respectively. The photospheric contribution is

$$R_{\rm phot} = -4.898 + 1.918(B-V)^2 - 2.893(B-V)^3 \tag{1.7}$$

and the total normalised flux is

$$R_{\rm HK} = 1.34 \times 10^{-4}\, C_{\rm cf}\, S_{\rm MW}, \tag{1.8}$$

where $S_{\rm MW}$ is the star's S-value calibrated to the Mount Wilson system and the conversion factor $C_{\rm cf}$ is defined as

$$\log C_{\rm cf} = 1.13(B-V)^3 - 3.91(B-V)^2 + 2.84(B-V) - 0.47. \tag{1.9}$$

Large stellar samples tend to show a distinctly bimodal $\log(R'_{\rm HK})$ distribution with active and inactive subsamples, separated by the "Vaughan-Preston gap" (Vaughan & Preston 1980; Gray et al. 2006). Stellar age-activity calibrations using $\log(R'_{\rm HK})$ have been constructed primarily via cluster stars with well-established ages (Mamajek & Hillenbrand 2008).

1.3.2 The basal limit of stellar activity

In the study of stellar activity, the Mount Wilson survey provided early evidence for a lower limit of chromospheric Ca II H & K emission (Schrijver 1987 and references therein). Plots of the chromospheric surface flux in several emis-

sion lines (including Ca II H & K and Mg II h & k) against stellar colour show clearly defined lower boundaries, with very few outliers below the lower limits. The minimum Ca II H & K line core flux, seen for the most inactive stars, has a chromospheric component that is independent of stellar cycle variability and referred to as "basal" emission. This means there are three contributions to measured S-values: the photospheric flux, the non-variable basal chromospheric component, and the magnetic activity related, time-variable chromospheric contribution (Schrijver et al. 1989).

The basal emission appears to be completely independent of large-scale stellar fields driven via the global magnetic dynamo. It was immediately suspected to arise from the dissipation of non-magnetic energy, via acoustic waves. This hypothesis is supported by detailed theoretical work (e.g. hydrodynamic modelling in Buchholz et al. 1998). However, there is still some debate and ongoing work on the exact mechanisms responsible for basal heating of the chromosphere (Hall 2005 and references therein). It now appears likely that both acoustic and small-scale magnetic mechanisms, unrelated to the stellar dynamo and its activity cycle, are at play (Hall 2005).

An important observational contribution was made by Schröder et al. (2012), who measured the solar S-index during the very pronounced activity minimum of 2008-2009. At several epochs, the solar disk was completely free from plages[9], and the overall emission from this "quiet sun" state was exactly at the basal limit observed in large stellar samples. This is evidence for the basal limit as a universal phenomenon for sun-like stars that is reached when they are devoid of active regions.

There have been a series of efforts to quantify the basal limit and investigate its dependencies on fundamental stellar parameters. Divergent methodologies are used, and results are quoted in terms of S-values, $\log(R'_{\rm HK})$ and $F'_{\rm HK}$ (e.g.

[9] the most relevant manifestation of active regions for chromospheric Ca II H & K emission

Isaacson & Fischer 2010, Wright 2004, Perez Martinez et al. 2014). The latter parameter is the chromospheric line core surface flux in absolute flux units, i.e. without normalisation to the total bolometric flux. The approaches used typically rely on empirical fits to the lower envelopes of the distribution of a given parameter, via large stellar samples (Isaacson & Fischer 2010; Mittag et al. 2013). Some works employ subtraction of best-fit synthetic spectra from observed stellar spectra (e.g. Takeda et al. 2012; Perez Martinez et al. 2014). This allows a more precise calculation of the full photospheric contribution via the synthetic spectra, and more robust use of absolute flux units. Due to this diversity of approaches, there is no unique agreed upon basal limit in the literature. Modern approaches do yield quite consistent results however (e.g. Perez Martinez et al. 2014).

For the purposes of this thesis the exact value of the basal limit (and which units it is given in) is of secondary importance. As will be discussed in Section 1.4.3, I am primarily interested in outliers significantly below the lower envelope of the distribution of chromospheric emission values. I will use large stellar samples to illustrate this lower envelope in Chapter 2, clearly revealing outliers below this empirical limit. It is convenient to use the most widely available metrics in the literature, i.e. S-values and $\log(R'_{\rm HK})$. The more physically meaningful, colour-independent metric is $\log(R'_{\rm HK})$, which can be converted from S-values where necessary. As shown in Section 1.3.1 the conversion only requires the widely available stellar $B-V$ colour index. The more precise methodology of calculating the basal limit of $F'_{\rm HK}$ in Perez Martinez et al. (2014) relies on the knowledge of precise stellar effective temperatures ($T_{\rm eff}$). These measurements are not widely available, significantly limiting any sample size. While $T_{\rm eff}$ can be estimated directly from $B-V$, the conversion is not precise without knowledge of stellar metallicity and stellar surface gravity (g_*), which again are not available for the large stellar samples of interest (c.f. Section 2.1).

A complication in the basal limits of S-values and $\log(R'_{\rm HK})$ is that both quantities show diverging distributions for the different stellar luminosity classes (see e.g. Wright 2004; Jenkins et al. 2008; Mittag et al. 2013 and Section 2.3). This is due to changing photospheric contributions in all S-index bandpasses as a function of g_*, and different magnetic activity behaviour as a star moves through its major evolutionary phases.

Main sequence, subgiant and giant stars should therefore be separated into different samples when analysing $\log(R'_{\rm HK})$ distributions. An accurate way to do this is placing stars with precise parallax measurements on the Hertzsprung Russell (HR) diagram (Wright 2004), as will be done in this thesis (Section 2.3). Wright (2004) provided the first analysis along these lines, and I will adopt a very similar methodology here. In essence, main sequence stars are defined as objects that have not evolved significantly above the average main sequence (for details, see Section 2.3).

Following Wright (2004), objects with $\log(R'_{\rm HK})$ < -5.1 are either evolved, or main sequence stars with extremely rare anomalously low activity levels. **The adopted main sequence basal limit is therefore** $\log(\boldsymbol{R'_{\rm HK}}) = -5.10$**.** More evolved stars often fall below this level. This exact value of the basal limit is somewhat arbitrary, but in practice provides a conservative, useful threshold between anomalous outliers and the pronounced lower envelope seen in large samples of main sequence stars. This will be illustrated in Section 2.4 for the stellar sample investigated in my thesis. Note that only 9 main sequence stars below this basal limit in a sample of thousands of objects were identified by Wright (2004). With such a low occurance rate, large stellar samples are clearly required when searching for these anomalous objects.

1.4 Stellar activity and exoplanets

Note that a shortened version of parts of this Section was published in Staab et al. (2017)

1.4.1 Activity in the context of planet searches

In planet searches, stellar activity is important as a source of nuisance signals. Active stars show elevated intrinsic variability in both their photometry and radial velocity timeseries (often referred to as "jitter"). Activity-related RV variability will be discussed in Section 4.2 in more detail. Several studies (e.g. Wright 2005a, Isaacson & Fischer 2010) have established empirical relationships between stellar activity and RV jitter, typically via $\log(R'_{\rm HK})$. Compared to preceding work, Hillenbrand et al. (2014) used a younger, more active stellar population for this purpose, estimating that for stars with $\log(R'_{\rm HK})$= -4.0, jitter levels are larger than ~200 $\mathrm{m\,s^{-1}}$ (Figure 1.5). The possible effects of the activity driven stellar high energy flux on planetary mass-loss and atmospheric evolution has been widely studied theoretically, see e.g. Lammer et al. (2003); Erkaev et al. (2007); Lopez et al. (2012); Koskinen et al. (2013); Jin et al. (2014) and Chadney et al. (2015).

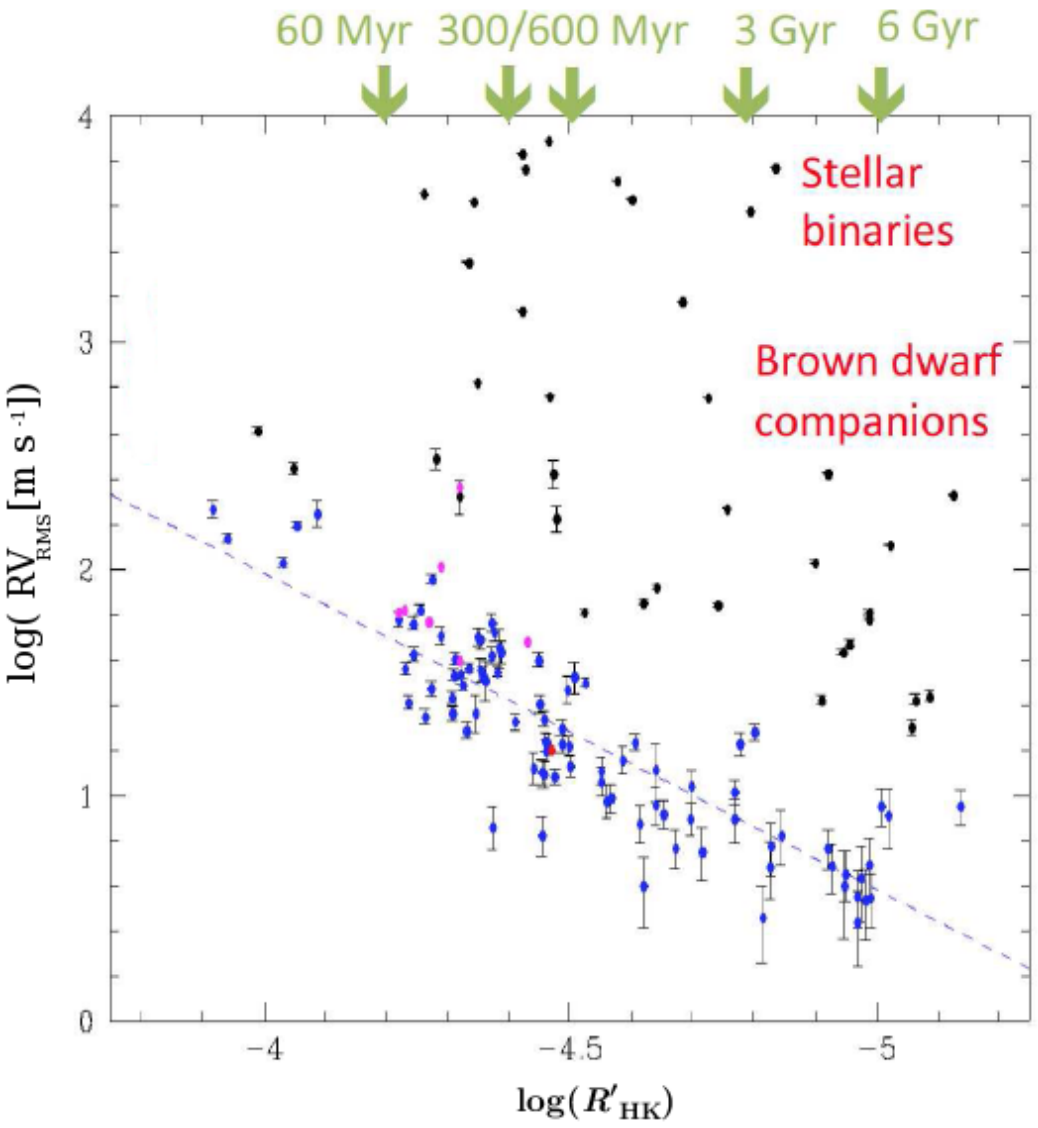

Figure 1.5: Figure adapted from Hillenbrand et al. (2014), showing their targets' observed radial velocity RMS against stellar activity level. Approximate ages corresponding to several $\log(R'_{\rm HK})$ levels are indicated. Hillenbrand et al. (2014) fit a linear relation (dashed) to 171 California Planet search targets (blue) after iterative sigma clipping of outliers (black). These are due to high-amplitude Keplerian signals from stellar and brown dwarf companions. Data from Lagrange et al. (2013) is shown for comparison in magenta.

1.4.2 Star-Planet Interactions

The hypothesis that close-in exoplanets can increase the stellar activity of their hosts through tidal or magnetic interaction was put forward by Cuntz et al. (2000). Such processes are often referred to in the literature as star-planet interactions (SPI). Stellar activity *enhancements* have been detected for individual HJ systems (e.g. Shkolnik et al. 2008; Pillitteri et al. 2011) and around periastron for a very eccentric system (Maggio et al. 2015). Null results for both chromospheric and coronal emission enhancements can be found in e.g. Poppenhaeger et al. (2010), Lenz et al. (2011) and Figueira et al. (2016). Activity *suppression* by a planet has been suggested by Miller et al. (2012) and Pillitteri et al. (2014) to explain the anomalously low X-ray and Ca II H&K line core emission of WASP-18. The parameters of this system, $M_P = 10.4\ M_J$, $P = 0.9$ d, suggest a particularly strong tidal interaction. The magnitude of this effect should be greater than for any other HJ system, and may significantly affect the shallow stellar convective zone of the F6V host, depressing its magnetic dynamo (Pillitteri et al. 2014).

Several studies have used large stellar populations to investigate whether mean stellar activity levels are systematically influenced by the presence and properties of planetary companions. Results have been very mixed for both X-ray measurements (e.g. Scharf 2010 in comparison with Poppenhaeger & Schmitt 2011) and Ca II H&K data (e.g. Canto Martins et al. 2011 in comparison with Krejčová & Budaj 2012). It seems that at best optical and X-ray diagnostics reveal magnetic SPI effects in only a small subset of observations of short period planets. Tidal spin-up of host stars by orbital decay of HJs could play a more important, but nonetheless limited role (Miller et al. 2015). The latter process would 'rejuvenate' the parent star, delaying or reversing the decline of rotation and activity with age. For a comprehensive overview of SPI studies and recent statistical work on this topic, see Miller et al. (2015). SPI detection

in large samples of systems remains controversial and beset by selection biases. However, strong evidence for activity enhancements in individual cases has been found in particular for planet-hosting wide binary systems (Poppenhaeger & Wolk 2014).

Theoretical work on magnetic SPI effects ranges from simple analytical approaches (e.g. Lanza 2008) to more sophisticated, three-dimensional MHD models (e.g. Strugarek et al. 2015). Studies such as Saur et al. (2013) and Lanza (2012) cannot account for the energy release observed for time-variable SPI signatures. Treatments which consider the dynamical behaviour of the magnetic field topology allow the phenomenon of magnetic reconnection to be included. Lanza (2012) finds that the power dissipated by reconnection between stellar and planetary fields at the planet's magnetospheric boundary is insufficient to explain observations attributed to SPI. In contrast, relaxation of stressed magnetic loops between stellar and planetary fields can provide sufficient power (Lanza 2013). Cohen et al. (2011) perform time-dependent MHD modelling, concluding that there is sufficient energy release from reconnection to explain observed SPI effects. These studies also provide explanations for the intermittent nature of magnetic SPI signatures seen in e.g. Shkolnik et al. (2008). However, both observational and theoretical work on SPI remain active areas of research without conclusive outcomes to date.

1.4.3 Circumstellar absorption of chromospheric emission

Haswell et al. (2012) first suggested that a high rate of planetary mass loss can feed diffuse circumstellar gas shrouds which absorb the stellar flux in the cores of strong resonance lines, depressing the chromospheric emission which arises in precisely these lines. In particular this naturally explains the anomalous zero flux Mg II h & k line cores observed in the extreme HJ host WASP-12. The line

core flux is zero[10], unique among all other stars of comparable age and spectral type. Extrinsic absorption across the entire stellar disc is strongly suggested by this, but cannot be explained by the interstellar medium along the line of sight, leaving the planet as a likely source of diffuse circumstellar gas.

The WASP-12 near-UV transit depth and duration observed by Haswell et al. (2012) imply exospheric gas is overfilling the Roche lobe and escaping. Note, however, that the hypotheses of mass loss from an exomoon (Ben-Jaffel & Ballester 2014) and Trojan satellites (Kislyakova et al. 2016) co-orbiting with WASP-12 have also been put forward. Corroborating the hypothesis of an enshrouded exoplanet system, Fossati et al. (2013) showed that WASP-12's Ca II H & K lines have an extremely low value of $\log(R'_{HK}) = -5.5$. If intrinsic to the star, this would be unique: WASP-12 is by far the lowest point shown in Figure 1.6. Occam's razor suggests these extreme properties of the star WASP-12 must be related to the presence of its extreme HJ planet, WASP-12b.

UV measurements of the Mg II h & k lines can only be made from space, and are therefore only available for a small number of stars. As discussed in Section 1.3, Ca II H & K observations can be readily carried out with ground-based facilities and are available for thousands of stars. Fossati et al. (2013) exploited this fact to show that five other HJ hosts, along with WASP-12, appear to show anomalously low $\log(R'_{HK})$ values compared to a large sample of stars (Figure 1.6). These objects are all well below the main sequence basal limit (Section 1.3.2). Note that Figure 1.6 also includes many subgiants below the basal limit.

The *apparent activity suppression* arising from planetary mass loss is sensitive to the planet's surface gravity. All else being equal, the lower the gravity, the higher the mass loss rate and the degree of circumstellar absorption.

[10] at all observed orbital phases

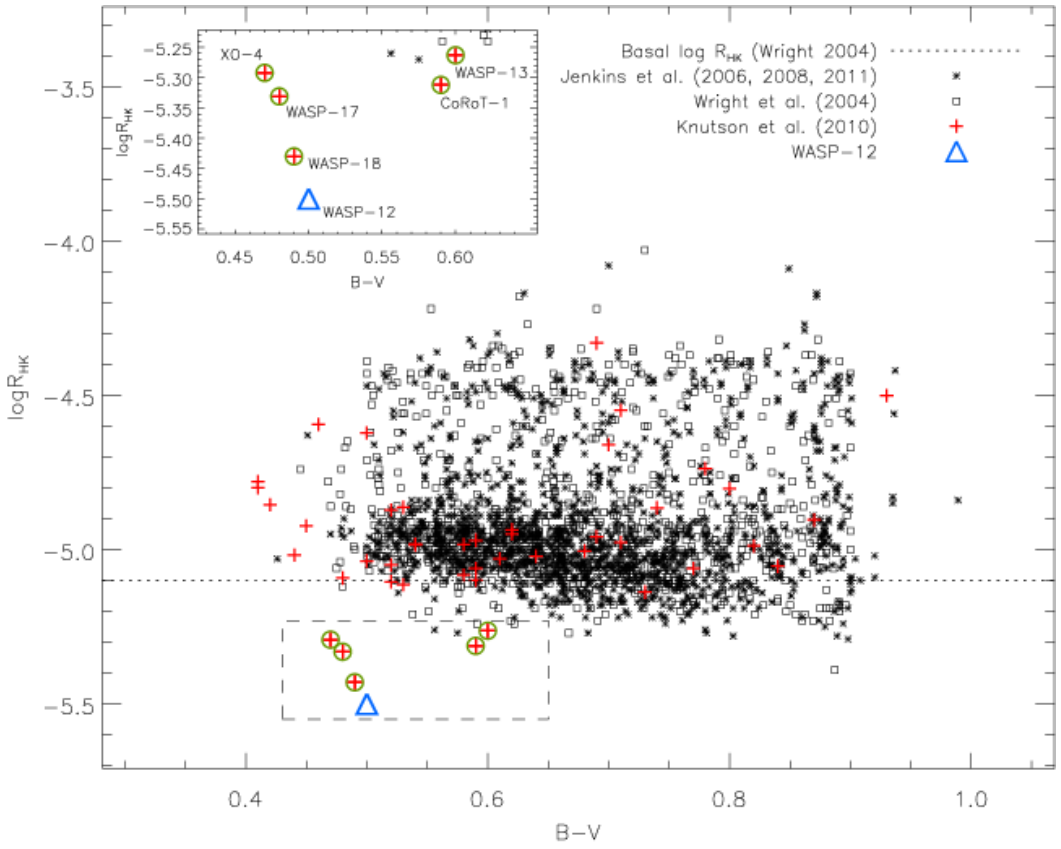

Figure 1.6: Figure taken from Fossati et al. (2013), showing the $\log(R'_{\rm HK})$ distribution of stars compiled from several catalogues and of the HJ hosts from Knutson et al. (2010). WASP-12 and five other outlying known exoplanet hosts are highlighted.

There is a highly significant correlation between planetary surface gravity ($g_{\rm p}$) and stellar $\log(R'_{\rm HK})$ for close-in planets[11] (Hartman 2010). Figueira et al. (2014) confirmed this correlation using a 3 times larger sample, and showed it is not due to selection biases. Lanza (2014) constructed a physical mass-loss model reproducing the observed correlation, which was refined by Fossati et al. (2015b). In the latter work, stars below and above the Vaughan-Preston gap, with their disparate *intrinsic* $\log(R'_{\rm HK})$ values, are treated as two distinct subsets. This approach may enable an estimate of the average effective stellar flux powering planetary mass loss, a major uncertainty in current planetary

[11] this was based on the Knutson et al. (2010) measurements for a sample of HJ hosts. Knutson et al. (2010) reported a correlation between stellar activity and the presence of a temperature inversion in the planetary atmospheres.

evolution models. However, a larger sample of planet hosts is needed; see details in Fossati et al. (2015b).

Systems enshrouded by material lost from close-in planets are highly promising targets for transmission spectroscopy, as discussed in Section 1.1.2. In the case of WASP-12, while the diffuse gas is present at all observed phases, it produces the most near-UV absorption close to transit (Haswell et al. 2012). Differencing the observed spectrum near transit and away from transit can reveal the additional absorption from the densest regions of the gas shroud, which can extend out to several planetary radii. Because the near-UV is such an informative wavelength region, with strong resonance lines of many abundant elements and ions, it is vital to make Hubble Space Telescope observations of the shrouded systems. There are currently no alternative means to obtain comparable information on planetary composition.

Large-scale planetary mass loss akin to that inferred from apparent activity suppression may underly the sub-Jovian desert (Kurokawa & Nakamoto 2014; Lundkvist et al. 2016). Other explanations of this dearth in the known exoplanet population have been suggested however (Matsakos & Königl 2016 and references therein).

It should be emphasised that systems with evidence for large-scale mass loss from Ly-α observations or anomalous $\log(R'_{\rm HK})$ values are currently distinct from systems with disintegrating rocky USPs detected via photometric dust cloud transits. The latter category is currently limited to very faint, distant stars where the former types of measurements are not feasible. Note also that planets in the former category do not necessarily have ultra-short orbital periods. Although WASP-12 and WASP-18 are in 1.1 and 0.9 day orbits, significantly longer periods are seen for other outliers highlighted in Figure 1.6 (P = 3.7-4.4 d); for GJ 436b (P = 2.6 d) and for 55 Cnc b (P = 14.7 d).

1.5 Thesis Overview

Interest in $\log(R'_{\rm HK})$ for individual known planet hosts is motivated by evidence it might provide regarding magnetic and tidal star-planet interactions and by the implications of anomalously low $\log(R'_{\rm HK})$ values for planetary mass loss processes. As a part of this thesis (Chapter 3), I therefore measure activity values of known hosts and assess the latest literature compilation of such values, searching for anomalous values in both cases. I also lay the groundwork for efficiently measuring the activity of a large number of known planet hosts by calibrating an instrument attached to a 10 metre class telescope to the Mount Wilson system.

My primary research question is whether stars without known planets and anomalously low $\log(R'_{\rm HK})$ values are in fact hosts of mass-losing, short-period planets. If this hypothesis (originated by my supervisor C.A. Haswell) is correct, it would provide a very efficient short-cut in the search for highly interesting, rare exoplanets (Section 1.1.2). It is worth emphasising that planets found by this novel approach would be virtually guaranteed to provide excellent opportunities for follow-up characterisation. Stellar activity catalogues consist of bright, nearby stars. Since planets would be found by virtue of a high mass loss rate, in-depth spectroscopic characterisation of the material lost should be possible. If the planets were to transit, this would entail transmission spectroscopy of the extended atmosphere. Even if there is no planetary transit, but the density of enshrouding gas along the line of sight from Earth varies in time[12], characterisation could be feasible. With current technological capabilities, the dearth of known mass-losing planets around bright stars (amenable to detailed characterisation) constrains the study of exoplanets. This means each discovery of such an object is important. If rocky USPs were to be discovered

[12] increasing at inferior conjunction

around bright stars, follow-up study of sublimated material could reveal the detailed surface composition of such planets, marking a step change in exoplanet science.

I therefore compile a target list of stars with anomalously low $\log(R'_{\rm HK})$ values in Chapter 2. I describe the planet search approach and methods used in Chapter 4, followed by a presentation of key results in Chapter 5. Final conclusions and suggestions for future work are given in Chapter 6. Where applicable, I detail the collaborative contributions to this thesis at the end of each Chapter.

Chapter 2

Archival Activity Data Compilation and Target Selection for Planet Search

In this Chapter, I will select an extensive stellar activity sample in Section 2.1, and search for anomalously low values. I discuss objects showing transient extreme activity outliers and possible causes of this in Section 2.2. In order to identify main sequence stars with activity values consistently below the basal activity limit (Section 2.4), I assess stellar evolution for my sample in Section 2.3. Additionally, targets in young stellar clusters are selected in Section 2.5. These are not below the basal limit, but show anomalously low activity in the context of their age and cluster membership.

2.1 Activity data compilation

As described in Section 1.3.2, a rigorous search for stars with anomalously low $\log(R'_{\rm HK})$ values requires a large compilation of measurements and an assessment of the evolutionary status of each star in the sample. Following the Mount Wilson survey, more than two dozen stellar activity catalogues have been published, with very heterogeneous properties. They use a range of spectrographs and differ dramatically in their sample size, number of measurements taken per star, and treatment of uncertainties. I provide an overview of this in Table 2.1. A robust and consistent statistical analysis of these catalogues is difficult: most provide no uncertainties for their activity measurements, and the number of measurements per star ranges over 3 orders of magnitude.

Table 2.1: Basic properties of the field star activity catalogues collected from the literature by Pace (2013). Note that many stars are listed in multiple catalogues: the total number of unique objects is 7864.

Publication	**Instrument/ Observatory**	**Sample size**	N_{obs} **per star**	**Comments**
Arriagada (2011)	Magellan Inamori Kyocera Echelle (MIKE)	673	1-127[a]	Magellan Planet Search Programme
Baliunas et al. (1995)	H-K photometers (HKP[b])	111	hundreds-thousands[c]	Mount Wilson survey (from 1966 to 1995).
Buccino & Mauas (2008)	International Ultraviolet Explorer (IUE)	259	1-306	From Mg II line emission converted to Ca II H & K scale.
Cincunegui et al. (2007)	Complejo Astronomico El Leoncito (CASLEO)	109	1-22	Study of connection between Ca II and H-α emission.
Duncan et al. (1991)	HKP	1296	1-1334[d]	Mount Wilson survey (from 1966 to 1983).

[a] only the mean or median values are published for each star
[b] the HKP-1 and HKP-2 photometers of the Mount Wilson survey are the only instances where no conventional spectrographs were used
[c] individual measurements are reported only in graphical format
[d] for each observing season, the maximum, minimum & mean values are reported

Table 2.1 – continued from previous page

Publication	**Instrument/ Observatory**	**Sample size**	N_{obs} **per star**	**Comments**
Gray et al. (2003)	Dark Sky Observatory (DSO)	664	1	Nearby Stars Project, Northern Sample. Non-linear calibration to Mount Wilson system; includes both low and medium resolution (R ~ 2000; 10 000) spectra from two different instrument setups.
Gray et al. (2006)	Cerro Tololo Inter-American Observatory (CTIO), Steward Observatory (SO)	1676	1	Nearby Stars Project, Southern Sample. Bootstrapped to non-linear calibration of Gray et al. (2003); includes low resolution (R ~ 1500) data.
Hall et al. (2007)	Solar-Stellar Spectrograph (SSS)	99	hundreds-thousands[c]	Long term survey successor to Mount Wilson Programme.
Henry et al. (1996)	CTIO	815	1-7	Characterisation of nearby Southern Hemisphere stars. Non-linear calibration to Mount Wilson; low resolution (R ~ 2400).

Table 2.1 – continued from previous page

Publication	**Instrument/ Observatory**	**Sample size**	N_{obs} **per star**	**Comments**
Isaacson & Fischer (2010)	High Resolution Echelle Spectrometer (HIRES), Hamilton Spectrometer	2630	1-1416	California Planet Search Programme
Jenkins et al. (2011)	Fiber-fed Extended Range Optical Spectrograph (FEROS)	890	1	Characterisation of bright stars in the solar vicinity.
Lockwood et al. (2007)	HKP	32	thousands[c]	Subset of Mount Wilson survey (from 1966 to 2004), focussed on a few stars with contemporaneous long-term photometric data.
Schröder et al. (2009)	FEROS, FOCES	481	1	Includes new method of activity measurements for rapid rotators.
Strassmeier et al. (2000)	Kitt Peak National Observatory (KPNO)	382	1-2	Only absolute H & K emission flux values given; transformed to S-values by Pace (2013).

Table 2.1 – continued from previous page

Publication	**Instrument/ Observatory**	**Sample size**	N_{obs} **per star**	**Comments**
López-Santiago et al. (2010)	Fiber Optics Cassegrain Echelle Spectrograph (FOCES), Intermediate Dispersion Spectrograph (IDS), Soviet Finnish High Resolution Echelle Spectrograph (SOFIN)	79	1-27	Characterisation of active young stars in the solar neighborhood.
Tinney et al. (2002)	University College London Echelle Spectrograph (UCLES)	59	1	Anglo-Australian Planet Search
White et al. (2007)	Palomar	390	1-7	Characterising nearby young stars, supporting Spitzer infrared survey.
Wright et al. (2004)	HIRES, Hamilton Spectrometer	1228	1-394	California and Carnegie Planet Search

The only large "catalog of catalogues" that had been published at the time of writing was Pace (2013), hereafter P13, containing 7864 stars. P13 dramatically simplified the underlying datasets by publishing only the minimum and maximum S-index ($S_{\rm min}$, $S_{\rm max}$) and $\log(R'_{\rm HKmin})$, $\log(R'_{\rm HKmax})$ values found among all published entries for each star. P13 then used the mean of these values ($S_{\rm mean}$, $\log(R'_{\rm HKmean})$) as an approximation of the mean stellar activity level. Obviously, more robust treatments would be possible, but are complicated by the datasets' heterogeneity. In addition, the majority of stars only have single-epoch measurements. Weighted means of all measurements for a given star cannot be obtained, since only a minority of measurements report uncertainties. The difference between the mean of all measurements and the mean of $S_{\rm min}$ & $S_{\rm max}$ values will be small for "well-behaved" stars, assuming the timeseries do not contain significant outliers. This is nicely illustrated by the very long baseline S-index timeseries of the Mount Wilson program, plotted in Figure 1 of Baliunas et al. (1995). Gomes da Silva et al. (2014) used a sample of 271 (mostly low-activity) stars, obtained over a $\sim$ 10 year timescale, to show that astrophysical variation in $\log(R'_{\rm HK})$ results in an RMS of typically 0.015 dex, and at most 0.08 dex. Very young, highly active stars can have significantly larger $\log(R'_{\rm HK})$ variation, at the $\sim$ 0.6 dex peak-to-peak level (see Figure 12 in Hernán-Obispo et al. 2015). Such objects are not relevant in the search for anomalous low activity values.

Ideally, weighted means of all measurements should be constructed, using uncertainties where they are reported, and accounting for small zero point offsets at the $\sim$ 0.1 dex level between different catalogues, as identified by Jenkins et al. (2011). Alternatively, a simpler option would be to use the median of all measurements, as this should also be more robust than the mean of $S_{\rm min}$ & $S_{\rm max}$ values. I opted to assess the P13 compilation of $S_{\rm min}$ and $S_{\rm max}$ values.

2.1.1 Cross-match with XHIP catalog

To convert the S-values of P13 to $\log(R'_{\rm HK})$, and to assess stellar evolutionary status, a reliable source for $B-V$ values and absolute magnitudes was required. I chose the Extended Hipparcos Compilation "XHIP" (Anderson & Francis 2012), which cross-references the final Hipparcos reduction output (van Leeuwen 2007) with several other sources for stellar data, including component designations of multiple star systems, cluster membership, spectral classification (validated against Hipparcos photometry) and metallicities. The two critical XHIP parameters of interest, $B-V$ values and absolute magnitudes (M_V), are always sourced from the original, reliable Hipparcos data, but the large amount of supporting data in XHIP is useful.

I used TOPCAT (Taylor 2005) to cross-match the P13 and XHIP compilations by stellar coordinates, using the closest match for each P13 entry. I discarded objects with no XHIP match (within 2 arcmin) and all objects that had an XHIP flag for combined photometry of stellar multiples. For these stars $B-V$ values may be unreliable, as flux from nearby companion(s) leaks into the photometric aperture. Finally, all remaining objects that did not have $B-V$ or M_V values were removed, leaving 5730 stars from the original 7864 entries in P13.

While searching for anomalous activity values, only the 5004 stars in the range $0.4 < B-V < 1.2$ were considered because $\log(R'_{\rm HK})$ is well calibrated in the range of 0.4 - 1.0 (Isaacson & Fischer 2010), and notably unreliable for $B-V > 1.2$ (Noyes et al. 1984).

2.1.2 A conversion problem in the Pace (2013) dataset

P13 compiled S-values from the available literature, and used these to re-calculate $\log(R'_{\rm HK})$ for a subset of his sample (2085 stars). Instead of using measured $B-V$ values for this calculation, P13 converted precision $T_{\rm eff}$ values from Casagrande et al. (2011) to $B-V$. I identified a significant problem with the P13 $\log(R'_{\rm HK})$ values: they showed a spurious disagreement with the values from the original catalog sources. I illustrate this in Figure 2.1, by showing only P13 objects with a single activity measurement compiled from Arriagada (2011). I crossmatched these P13 entries with the original Arriagada (2011) catalog. The S-values from both agree exactly, while $\log(R'_{\rm HK})$ differs typically at the 0.15 dex level, and up to 0.6 dex in one case (Figure 2.1). The disagreement $\Delta\log(R'_{\rm HK}) = \log(R'_{\rm HK})_{\rm P13} - \log(R'_{\rm HK})_{\rm Arriagada}$ is dominated by a systematic effect as a function of $T_{\rm eff}$. Similar behaviour was seen for other source catalogs. I verified that my $\log(R'_{\rm HK})$ values (calculated from the P13 S-values and $B-V$ from XHIP) agree with the original catalogs. This confirms that the problem lies in P13's implementation of the Noyes et al. (1984) conversion and/or his conversion from $T_{\rm eff}$ to $B-V$. With insufficient details on these steps given in P13, the exact cause is uncertain, but may simply be a typo in one of the $B-V$ coefficients of Equations 1.7 or 1.9 (Section 1.3) as used by P13. The issue does not affect the work detailed in this thesis, as I only used the S-values compiled by P13.

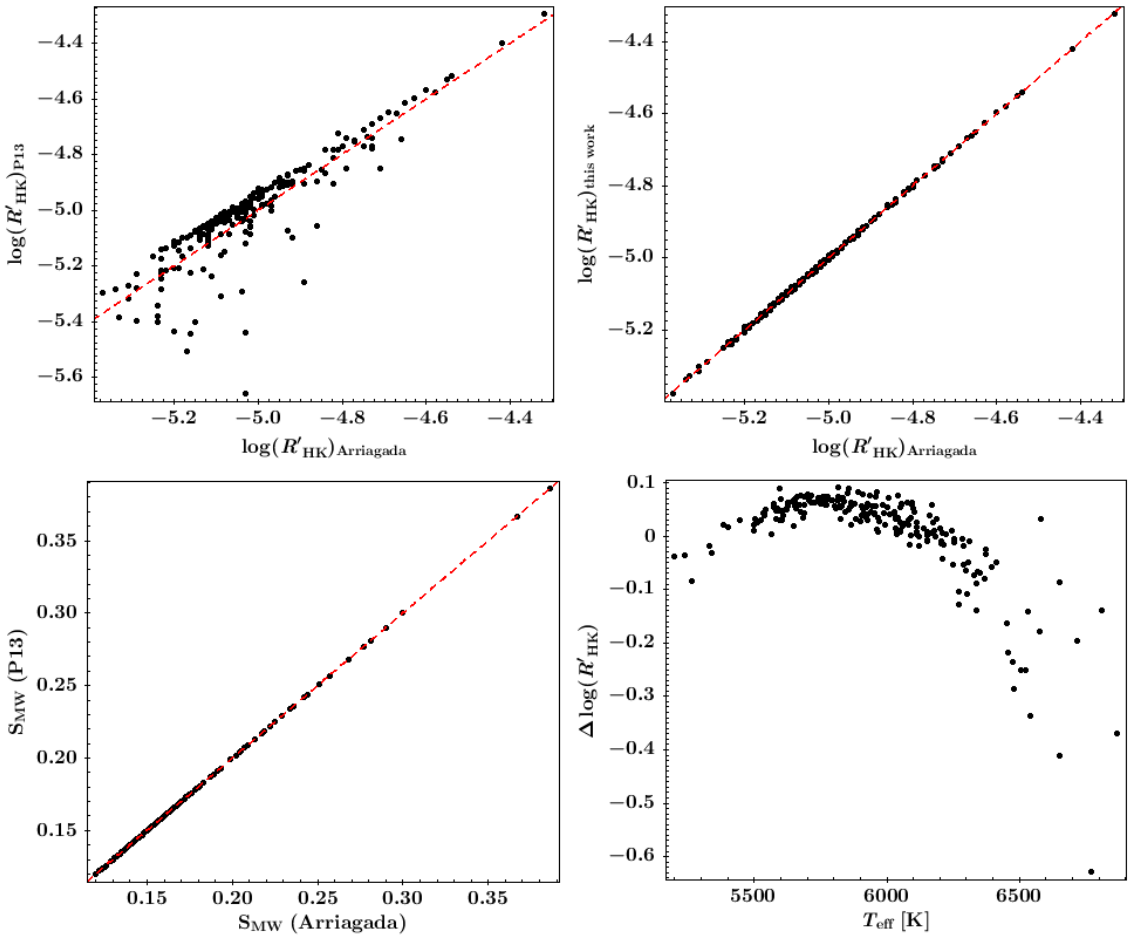

Figure 2.1: There are large discrepancies between P13 and Arriagada (2011) $\log(R'_{\rm HK})$ values (**top left**), although they are based on identical S-values in both catalogues (**bottom left**). The dashed lines mark the 1:1 relationships. My $\log(R'_{\rm HK})$ values (**top right**), calculated from P13 S-values and $B-V$ from XHIP, agree with those in Arriagada (2011). The difference $\Delta\log(R'_{\rm HK}) = \log(R'_{\rm HK})_{\rm P13} - \log(R'_{\rm HK})_{\rm Arriagada}$ is shown as a function of stellar effective temperature (**bottom right**).

2.2 Activity "dropouts"

Figure 2.2 shows the $S_{\rm min}$ and $S_{\rm max}$ distributions of the P13 sample, compared to the S-values of the HJ hosts reported in Knutson et al. (2010).

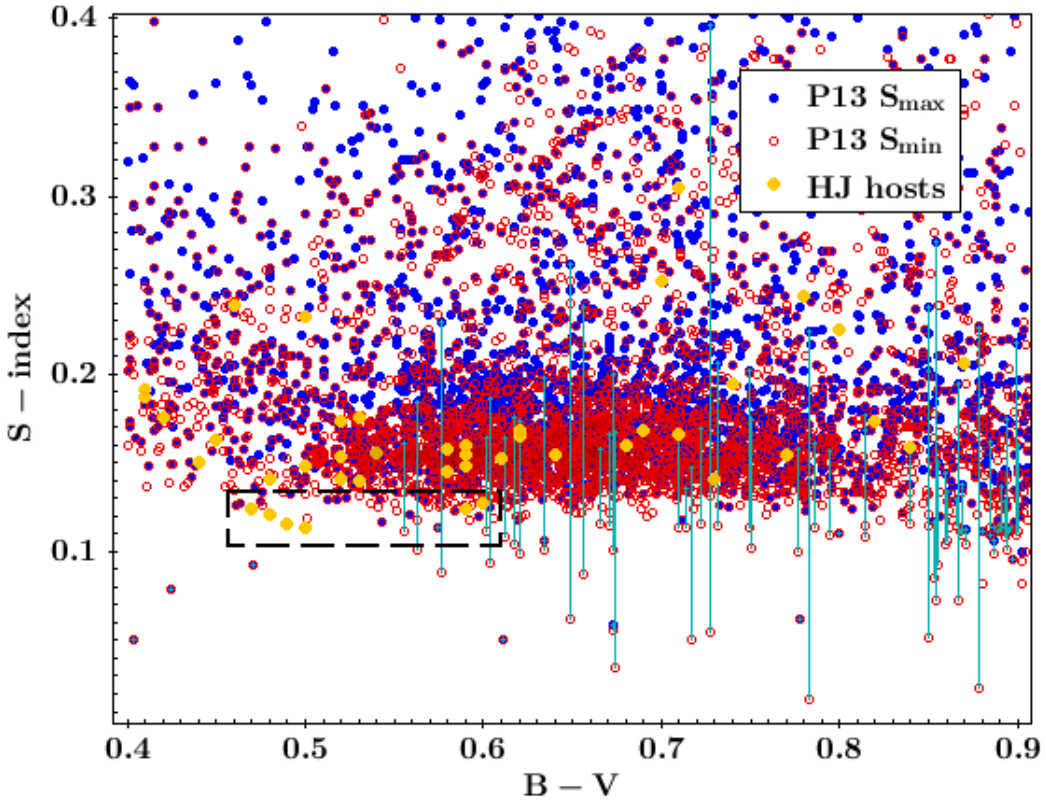

Figure 2.2: $S_{\rm min}$ and $S_{\rm max}$ distributions of the P13 sample and activity values of the HJ hosts in Knutson et al. (2010). $S_{\rm min}$ values significantly below the basal envelope of the distribution are joined to the corresponding $S_{\rm max}$ values with vertical lines. The HJ outliers identified by Fossati et al. (2013), c.f. Figure 1.6 (Section 1.4.3), are highlighted.

There are a number of conspicuously low activity outliers well below the basal envelope and below the known HJ outliers, several of which have $S_{\rm max}$ values within the main above-basal distribution. I refer to these as activity "dropouts". They include extremely low minimum S-values ($\lesssim 0.1$), where $\log(R'_{\rm HK})$ becomes undefined. The line core flux reported for these cases is lower than the photospheric flux contribution, $R_{\rm phot}$, leading to a negative $R_{\rm HK}$. If these values have a genuine astrophysical origin, they are prime candidates for extrinsic absorption signals and cannot be intrinsic to the star. There are

no known stellar phenomena that would depress the intrinsic emission below not only the basal chromospheric level, but below the photospheric flux level. Objects with such highly divergent apparent $S_{\rm min}$ and $S_{\rm max}$ values (at the 30 - 90% level) are an unexpected result, and warrant close examination. There are possible systematic effects that can depress S-values. Conceivable options are uncorrected cosmic ray hits or contaminating light in the continuum V and R channels, or background over-subtraction in the data reduction. The latter has a small effect on the V & R bandpasses but would depress the line core flux appreciably. The work in Fossati et al. (2017) showed that an over-subtraction at the 5% level for a star near the basal limit ($S \sim 0.14$) leads to a depression of the measured S-value by 50%. Note that a lack (or under-performance) of background subtraction has the opposite effect, leading to spuriously high S-values. Contaminating light is known to cause problems for measurements of S-values from HARPS spectra and has to be corrected for (Lovis et al. 2011). Ruling out such systematic effects is possible with full spectral analysis, including re-reduction with careful background subtraction routines. Unfortunately, spectra from which the activity measurements in P13 were derived are usually not publicly available. In the case of the spectrophotometric Duncan et al. (1991) Mount Wilson data, spectra were in fact never recorded.

I found 6 conspicuous stars that have $S_{\rm max} = S_{\rm min}$ values below the photospheric contribution. These are the entries in Figure 2.2 with S-values $\lesssim 0.1$ without vertical lines. They are all extremely evolved stars (supergiants). All but 1 of them have only a single-epoch Mount Wilson (Duncan et al. 1991) measurement. Their reliability can therefore not be assessed through spectroscopic analysis, and new spectroscopic measurements are needed to re-assess their S-values. Circumstellar absorption from massive stellar mass-loss in this evolutionary phase could be the astrophysical cause of these outliers, if they are real. Note that these stars were later removed from the sample as my work focussed on main sequence stars (see Section 2.3).

I found that the majority of "dropout" stars with S_{max} above and S_{min} values well below the basal limit arise from data taken with the Lick observatory's Hamilton Spectrograph, reported by Isaacson & Fischer (2010). The earlier Wright (2004) publication of a subset of this data conservatively discarded all Lick spectra with SNR < 50, to avoid unreliable values, while Isaacson & Fischer (2010) reported Lick values regardless of SNR. I found that the low SNR regime can cause outliers that are very likely to be spurious. Systematic issues as discussed above probably come into play in addition to photon noise scatter. It was not possible to assess this further, as Lick spectra are not public and Isaacson & Fischer (2010) do not report uncertainties for their activity values (providing only the spectral SNR in the continuum). I did not find a unique SNR threshold below which a dramatic increase of activity scatter occurs. This is not unexpected, since the ratio between SNR in the continuum and in the line cores can vary substantially, depending on the spectral type and mean activity level.

I also identified instances where Lick data reported high SNR, but nevertheless showed spurious activity measurements. As shown in Figure 2.3, several stars have extreme dropouts within the same week of observations. The only convincing explanation for this is an instrumental or data reduction artefact. Since P13 reports S_{max} & S_{min} values from the entire literature for every star, the few outliers in multi-epoch timeseries of Isaacson & Fischer (2010) were propagated into the P13 sample. In view of these spurious outliers affecting $\log(R'_{HKmin})$ and therefore $\log(R'_{HKmean})$ in Pace (2013), along with his incorrect conversion of all S-values as identified in Section 2.1.2, the conclusions of that paper should be revisited.

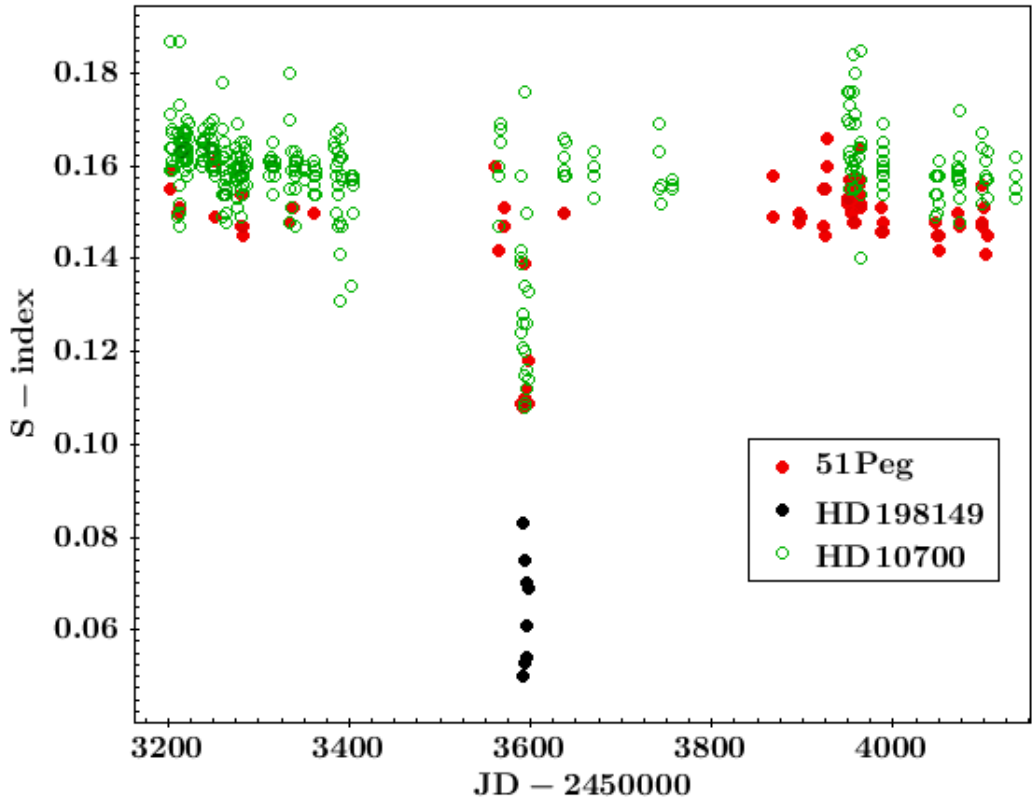

Figure 2.3: Spurious activity dropouts in a subset of the Lick (Isaacson & Fischer 2010) timeseries of 3 stars all occur in the same week (2453589 < JD < 2453597). For 51 Peg and HD 10700, the dropouts are from their normal activity level to the sub-basal regime, while all Lick measurements for HD 198149 are spurious and appear below the photospheric level. An instrumental or data reduction issue must be the cause of such anomalies.

2.2.1 Dropouts caused by transit events?

I could not attribute all dropouts found in the P13 sample unambigously to the instrumental or SNR problems described above. Astrophysical phenomena causing such signals are limited. Since the stellar basal limit corresponds to a completely quiet state (Section 1.3.2), dramatic, transient drops of the intrinsic emission below this level are not expected. If extrinsic absorption is the cause of these changes on timescales of hours to days, it must arise from highly time-variable gas in the stellar systems themselves rather than in the interstellar

medium (ISM). For the stars in the P13 sample, typical proper motions will move our line of sight relative to the ISM by on the order of 0.01 AU d^{-1}. The local interstellar medium has not been observed to vary on these very small scales (Montgomery & Welsh 2012).

Planetary transits

To my knowledge, there has been no secure detection of Ca II in an exoplanetary atmosphere to date. Mura et al. (2011) modelled planetary mass-loss from the USP CoRoT-7b and predicted a relatively high escape rate of Ca II, causing a highly elongated (up to 2 R_*) comet-like tail, that should be detectable in transit. However, Guenther et al. (2011) then presented a null-result for Ca II and several other species from UVES observations of the transit of CoRoT-7b, using long exposures with high SNR. Ridden-Harper et al. (2016) conducted an in-depth spectroscopic search for Na and Ca II in the transit of 55 Cnc e, with a multi-instrument campaign covering 5 transits. 55 Cnc is a very bright (V = 6.0) host of a multi-planet system (Fischer et al. 2008) that includes a warm Jupiter and the transiting USP super-Earth 55 Cnc e. This planet has been studied extensively with several high profile results (e.g. Winn et al. 2011, Madhusudhan et al. 2012, Demory et al. 2016, Demory et al. 2016), though the nature of its atmosphere, surface, and composition remains enigmatic (Ridden-Harper et al. 2016 and references therein). A significant (4.9 σ) Ca II signal was detected in the HARPS data of only one of the transits, with null results for the other four. As a result, Ridden-Harper et al. (2016) do not formally claim the detection, but if real it would correspond to Ca II gas from a highly variable exosphere, beyond the planet's Roche-Lobe.

Since no activity values are provided in Ridden-Harper et al. (2016), I used

the TERRA package (Anglada-Escudé & Butler 2012; Section 4.1.3) to extract S-values from the publicly available HARPS spectra covering the relevant transit. The resulting timeseries is shown in Figure 2.4. It illustrates that the possible Ca II transit event does leave a clear signal in the S-values, despite their bandpasses not being optimised for spectroscopic transit detection. During mid-transit, the S-value is depressed by ~ 1.5 %, corresponding to a reduction in $\log(R'_{\rm HK})$ by ~ 0.01 dex.

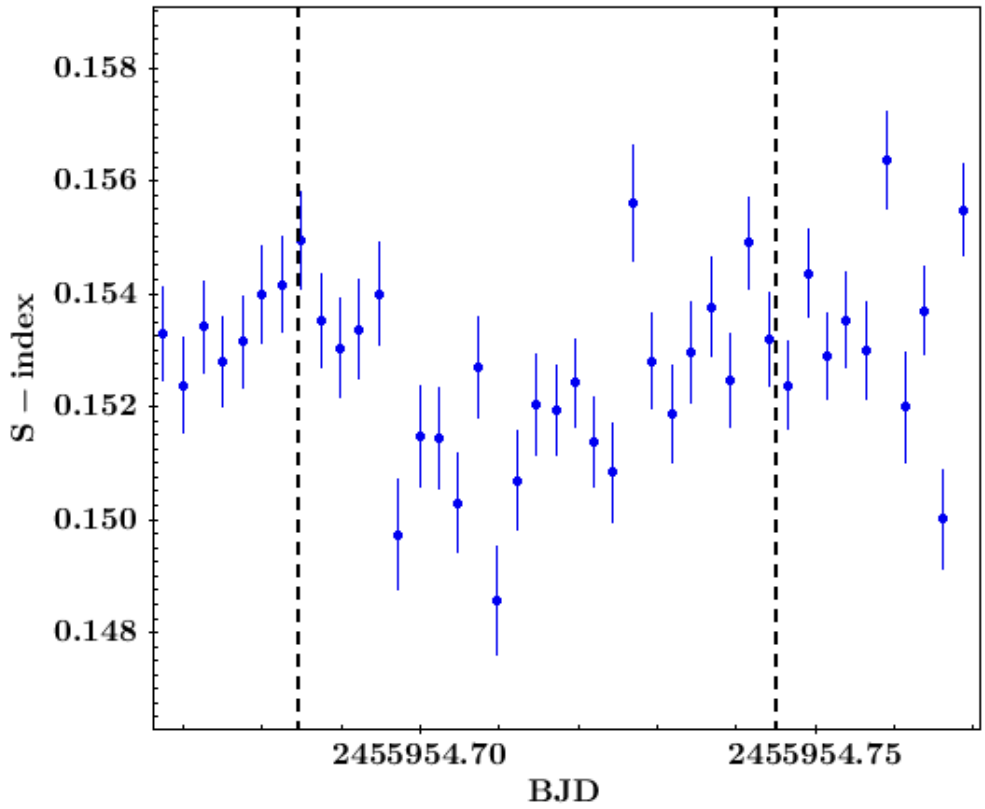

Figure 2.4: Possible Ca II transit of 55 Cnc e in the HARPS dataset analysed by Ridden-Harper et al. (2016). Start and end points of the solid-body planetary transit are marked.

Obviously this is a much smaller signal than the activity dropouts discussed in Section 2.2. 55 Cnc is a relatively inactive star (Fischer et al. 2008). The large, high-quality HIRES dataset published by Isaacson & Fischer (2010) nevertheless shows intrinsic peak-to-peak $\log(R'_{\rm HK})$ variability of 0.05 dex due to stellar rotation on timescales of days. The rotation period of 55 Cnc is ~43 days, as

detected in both photometric and $\log(R'_{\rm HK})$ timeseries (Fischer et al. 2008). On timescales of years, the Isaacson & Fischer (2010) dataset shows peak-to-peak $\log(R'_{\rm HK})$ variations of 0.15 dex, due to the stellar activity cycle. The RMS over the $\sim$ 5 yr baseline is 0.05 dex, and no dropouts comparable to those in Section 2.2 are present. Ridden-Harper et al. (2016) present evidence that the possible transit signal does not arise from stellar activity variability on timescales of hours. However, without prior knowledge of the ephemeris a Ca II transit detection at the 0.01 dex level would certainly be extremely challenging.

Note that to date both CoRoT-7b and 55 Cnc e have not been shown to undergo significant mass-loss for any species. (Ehrenreich et al. (2012) presented a null result in the search for Lyman-α transits of 55 Cnc e, while detecting a possible transit of an extended exosphere around 55 Cnc b.) Before the possible 55 Cnc e Ca II transit detection had been announced, I explored the potential of HJ transits to cause transient activity dropouts through Ca II absorption. Of particular interest are HJ systems with published evidence of significant planetary mass-loss, including anomalously low $\log(R'_{\rm HK})$ values. There are publicly available HARPS spectra that cover HJ transits from several programs focussed on detecting the Rossiter-McLaughlin effect (e.g. Triaud et al. 2010). Unfortunately, the faintness of typical HJ hosts (V$\gtrsim$11), combined with the exposure times used during transit ($\sim$10 min), results in very poor SNR in the wavelength region of interest. We found that these activity timeseries are dominated by systematic effects, and not suitable for the detection of Ca II signals.

An exception to this rule is the high SNR HARPS dataset of the bright (V = 7.6) HJ host HD189733. This is by far one of the best studied exoplanetary systems, with many papers focussed on mass loss and spectroscopic transit detection. Escaping hydrogen has been detected repeatedly beyond the planet's Roche-

Lobe using space-based UV observations (Lecavelier des Etangs et al. 2010, Lecavelier des Etangs et al. 2012). Detection of sodium in the planet's atmosphere was accomplished in Redfield & Linsky (2008), the first ground-based detection of Na for any exoplanet. Cauley et al. (2015) report the detection of Hydrogen Balmer line absorption both in- and pre-transit, and interpret this as the signature of a bow-shock. The HARPS data assessed here were used to detect sodium by Wyttenbach et al. (2015), who present blueshifted absorption features potentially related to upper-atmospheric winds. None of these works searched for Ca II signals. The public HARPS data cover 3 transits, one of them partial. We found that one of the transits is strongly affected by a stellar flare, causing a spike in Ca II emission. Note the discovery paper of the planet (Bouchy et al. 2005) showed that the star is active. It has prominent Ca II emission cores and during the HARPS transits ranged between -4.55 $\leq \log R'_{HK} \leq -4.50$. I initially used S-values extracted via TERRA to identify a candidate Ca II signal in the partial transit. The S-values are depressed by $\sim$ 1.5% during mid-transit ($\sim$ 0.01 dex in $\log(R'_{\rm HK})$), as for the candidate 55 Cnc signal.

We then carried out an in-depth analysis for all 3 transits, including the behaviour of the NaD and Hα lines, and assessed if the apparent absorption features arise purely from activity variation or from genuine planetary absorption. The results are presented in Barnes et al. (2016). To summarise this work, we confirmed high-significance apparent absorption signals in the NaD and Hα lines, and reported the first such signal for the Ca II lines (at 1.7 and 9.8σ) in the 2 transits not affected by strong flaring. The apparent absorption depth in the Ca II and Hα lines shows significant variability from transit to transit. Time-resolved analysis of the line-profiles revealed that the signals in all lines do not follow the velocity-dependence expected from the in-transit planetary orbital motion. Instead, the signals are better recovered in the stellar frame of reference. Variable chromospheric contributions to all three absorption line

cores may therefore be the cause of the apparent absorption and its changing depth. Temporal variability of the chromosphere itself can cause systematic effects, along with the planet transiting across regions with divergent activity on the stellar disk. While a genuine planetary absorption contribution could not be ruled out, the detection of such signals in lines with a chromospheric contribution is clearly problematic. Our findings cast doubt on the purely planetary interpretation of Hα and NaD signals published by Cauley et al. (2015) and Wyttenbach et al. (2015) (see previous paragraph). More recently, Cauley et al. (2017) published an analysis of the Hα signals in seven HIRES and HARPS transits. They simulate the effect of the planet transiting across active regions in detail, and conclude that for specific model input parameters this can reproduce the apparent absorption signals observed. However, large faculae/plage emission strenths and coverage fractions, concentrated in the stellar latitudes traversed by the planet, are needed. The authors favor a planetary interpretation, as their models are able to explain the observed velocities of the Hα absorption, and since the stellar explanation requires very specific model constraints.

Exocomet transits

The young planet host β Pictoris was shown by Hobbs et al. (1985) to have strong circumstellar absorption in the Ca II lines associated with its gas and dust disk, but additional, Doppler-shifted transient absorption events were detected in many subsequent observations (Ferlet et al. 1987). These are entirely consistent with evaporating, stargrazing exocomets and have been extensively characterised and modelled in this context (e.g. Beust et al. 1998, Kiefer et al. 2014a). Cometary tails are able to cover large fractions of the stellar disk, up to 100%, causing very significant absorption in strong resonance lines. In the case of β Pictoris, simultaneous exocomet transits often result in multiple distinct absorption features in the star's Ca II line profiles. Figure 2.5 shows a typical β

Pictoris spectrum to illustrate this.

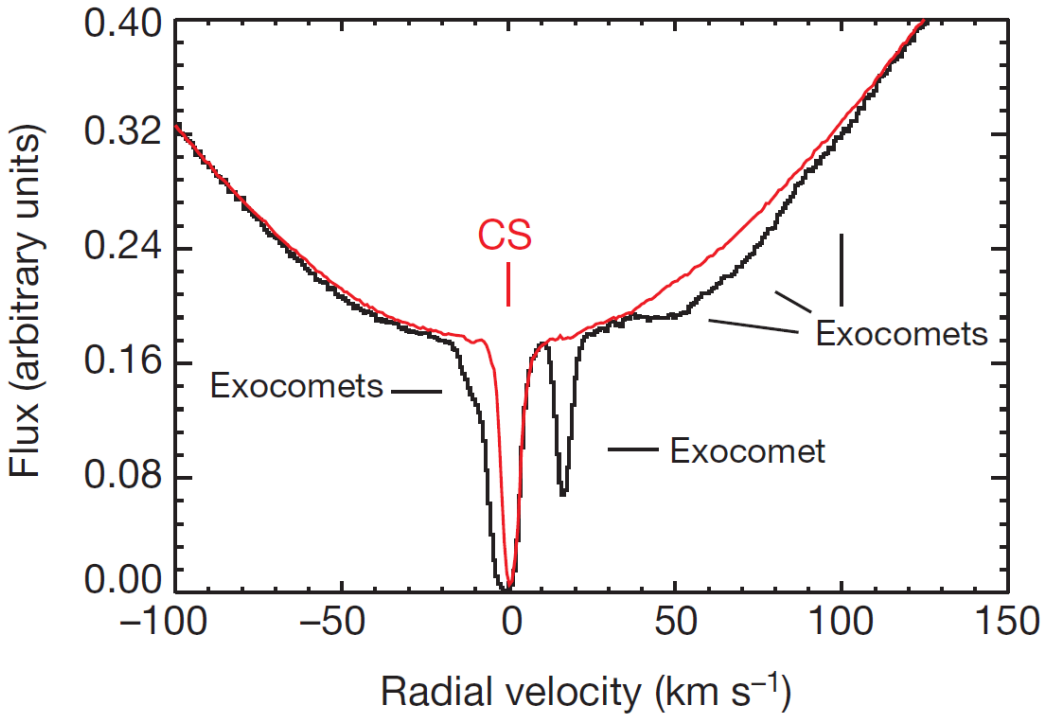

Figure 2.5: **Red:** The average, velocity-broadened Ca II K line profile of β Pictoris includes a sharp, stable circumstellar (CS) component. **Black:** A typical observation reveals multiple transient exocomet absorption components (some of which are blended), at different velocities. Figure taken from Kiefer et al. (2014a).

Exocomet absorption signals can occur within the *H* and *K* core bandpass windows if their velocity is between -86 and 86 km s^{-1}. The resulting change in the measured *S*-value should be quantified along the lines of the ISM absorption analysis done in Fossati et al. (2017). Qualitatively, a comparison of Figure 2.5 and e.g. Figure 2 of Fossati et al. (2017) makes it clear that the effect of such exocomet signals is non-negligible. Apart from β Pictoris, exocomet detections have been made for several other A-type stars (Montgomery & Welsh 2012, Kiefer et al. 2014b). These systems have very different properties than the vast majority of those in the P13 sample. They are all very young (<60 Myr), with suspected debris disks (Welsh & Montgomery 2013). Since they are A-type stars, they have no chromospheric Ca II H & K emission, and $\log(R'_{\rm HK})$ is not defined. Exocomet signals are relatively straightforward

to detect in their velocity-broadened Ca II lines (Fig. 2.5) without the complications of activity variability. Typical (much older, debris-free) F, G & K systems in the activity catalogues should have a much lower rate of exocomet infall. Nevertheless, for objects observed at many epochs over several years, the detectability of such events via low $\log(R'_{\rm HK})$ anomalies is an intriguing possibility. For our mature Solar System, a high rate (~270 $\rm yr^{-1}$, Knight 2008) of sungrazing comets has been revealed by SOHO observations. Some of these may cause detectable Ca II H & K transit events in the solar spectrum, particularly since their ephemerides are known. This line of investigation could be pursued with the ongoing very high-cadence, long-baseline HARPS-N solar observations (Dumusque et al. 2015), which cover Ca II at high SNR.

2.2.2 Avoiding dropouts in the search for activity anomalies

The cause of at least the vast majority of activity dropouts is unlikely to be mass-loss from short-period planets. I therefore used the $\log(R'_{\rm HKmax})$ values calculated from the P13 $S_{\rm max}$ entries in the search for consistently anomalously low activity levels. This avoids spurious entries of $\log(R'_{\rm HKmin})$ which also affect $\log(R'_{\rm HKmean})$. Such a conservative approach could miss stars of interest, where a few high activity values in their timeseries would prevent the detection of an anomalously low $\log(R'_{\rm HKmean})$. However, as discussed in Section 3.3.3 typical astrophysical variability is low (RMS of $\sim$ 0.015 dex), and many stars in the sample only have single-epoch measurements. In both cases, an anomalous activity level would be picked up by using a $\log(R'_{\rm HKmax})$ criterion. The benefit of the criterion I adopted is that the targets identified have a high probability of showing genuinely depressed Ca II H & K line core fluxes.

2.3 Stellar evolutionary status

Figure 2.6 illustrates the effect of post main sequence (MS) stellar evolution on the activity level. As noted previously, main sequence, subgiant and giant stars have very different activity distributions (see also Mittag et al. 2013). For

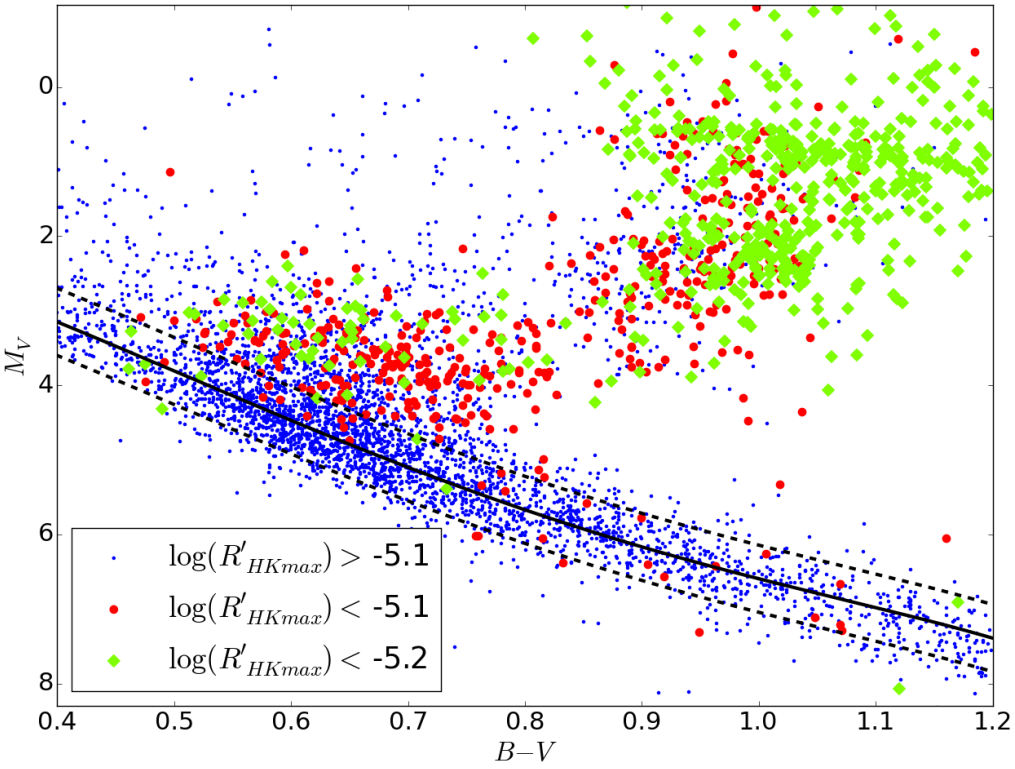

Figure 2.6: HR diagram of the 5004 P13 stars in the range $0.4 < B-V < 1.2$ with reliable XHIP data. The empirical average MS from Wright 2005b with the 0.45 mag range (dashed) is shown, corresponding to unevolved stars with approximately -0.3 < [Fe/H] > +0.3 (Wright 2004). My MS sample is confined to stars below the upper dashed line.

the work in this thesis, only unevolved main sequence stars were considered as potential targets. Extending the search for anomalous activity levels to subgiant and potentially giant stars is left to future work. The basal activity levels of these evolved objects are less well understood than for MS stars, and the complication of increasing stellar mass loss rates modifying the Ca II H & K line cores should be taken into account (c.f. Shetrone et al. 2008). The analysis of the activity levels of subgiants and giants done by Mittag et al. (2013) will

be a useful starting point for this work.

To define unevolved stars, I used the XHIP $B-V$ and M_V values and the Wright (2004) empirical average main sequence, $M_{V,MS}(B-V)$, to calculate the distance from this MS for each star, ΔM_V:

$$\Delta M_V = M_{V,MS}(B-V) - M_V \tag{2.1}$$

In the correct version from Wright (2005b), the MS is defined as a ninth-order polynomial expression: $M_{V,MS}(B-V) = \sum a_i (B-V)^i$, with coefficients a = {1.11255, 5.79062, -16.76829, 76.47777, -140.08488, 127.38044, -49.71805, -8.24265, 14.07945, -3.43155}. This average MS is shown in Figure 2.6.

As in Wright (2004), an unevolved main sequence population is defined as $\Delta M_V < 0.45$, which retains stars up to a metallicity of [Fe/H] $\sim$ +0.3. This is a somewhat arbitrary choice, and represents a compromise between retaining metal-rich, unevolved stars, while rejecting (metal-poor) evolved stars. **2716 out of my** 5004 **P13 stars** in the range $0.4 < B-V < 1.2$ **satisfy this definition of being unevolved, and constitute my final primary sample in the search for anomalously low activity values.** Figure 2.7 shows the visual magnitude and distance distributions of this sample, **referred to as "MS sample" hereafter**, compared to the known exoplanet host population. Similarly to the non-transiting planet hosts in the literature, the MS sample has distances below $\sim$ 100 pc and $5 \lesssim V \lesssim 11$. This is unsurprising, as many of the catalogs compiled by P13 are published in connection with RV surveys. The MS sample is therefore by design well-suited to RV planet searches and detailed follow-up of any discoveries: it contains bright, nearby stars. This stands in contrast to the transit host population, which is biased towards significantly fainter, more distant systems. Known ultra-short period planets and/or planets thought to be undergoing significant mass-loss are mostly found in this transit host

population (by the Kepler survey, see Chapter 1). This makes discoveries of such planets in my MS sample particularly valuable cases for follow-up studies. Note the $B-V$ selection criterion limits the MS sample to stars of spectral type $\sim$ F4 - K5 (c.f. Pecaut & Mamajek 2013). Similarly to the known exoplanet host population, it is dominated by late F-type and G-type stars.

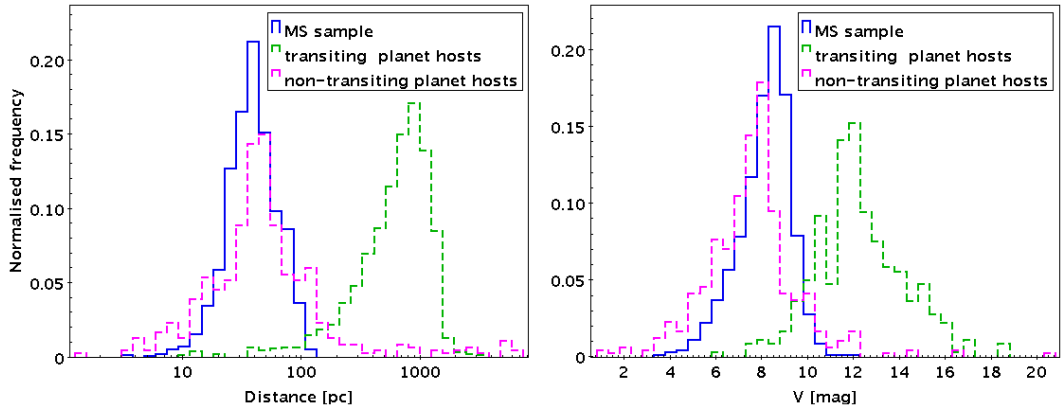

Figure 2.7: Magnitude and distance distributions of my MS sample compared to known planet hosts from the Exoplanet Orbit Database.

2.4 Targets below the basal limit

Figure 2.8 shows the activity - colour distribution of my MS sample, using $\log(R'_{\rm HKmax})$. For a full, updated comparison with the known short-period planet distribution, see Section 3.4.1. The minimum requirement for a star to be classified as a target of interest was $\log(R'_{\rm HKmax}) < -5.10$, i.e. maximum measured activity level below the adopted basal limit. There are 39 of these objects in the MS sample, which became the primary targets of interest for our planet search programme. Note that the majority of stars in the MS sample and of these targets have single-epoch activity measurements.

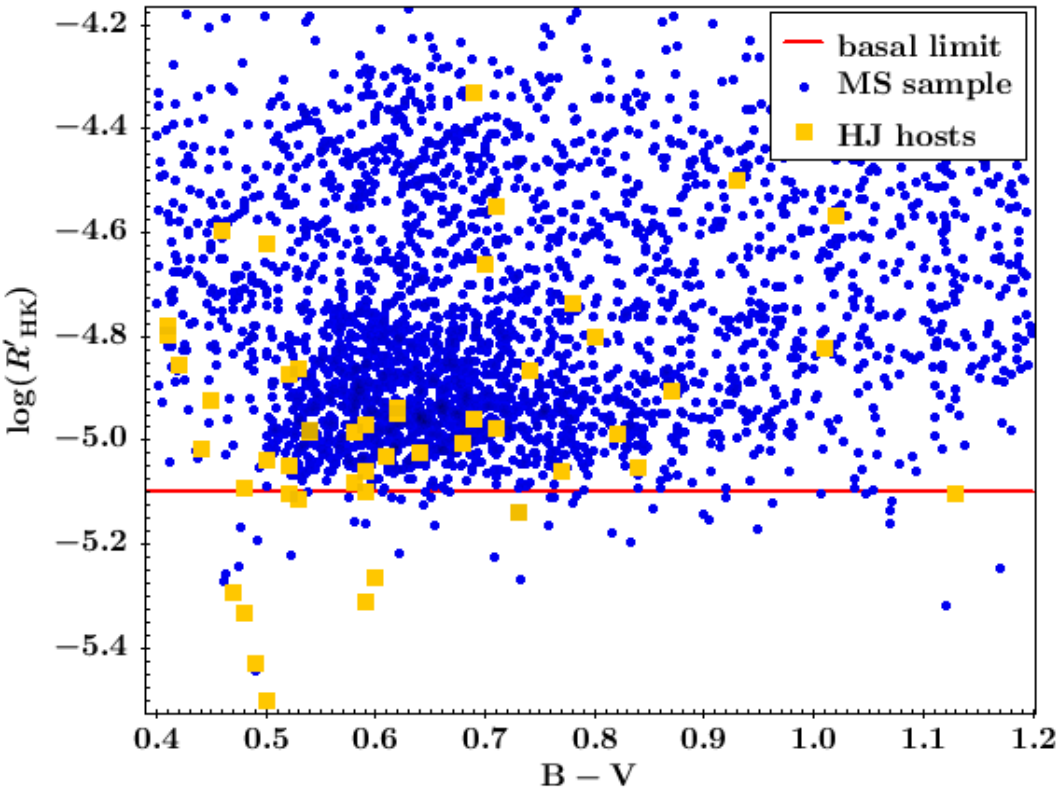

Figure 2.8: $\log(R'_{\rm HKmax})$ distribution of my MS sample compared to known HJ hosts from Knutson et al. (2010).

In addition to the $\log(R'_{\rm HKmax})$ requirement, practical observational constraints influenced the prioritisation of targets in our RV planet search. As a consequence of the relevant telescope time allocation rules and pressures, and the required high cadence (see Section 4.4) it was clear that RV observations

would be primarily carried out in dedicated visitor-mode runs over ~ 3-6 consecutive nights. When applying for telescope time in a given semester and hemisphere, I allocated the highest priority to targets with lowest $\log(R'_{\rm HKmax})$, grouped by "required" coordinates. Stars with archival RV measurements showing significant variability, particularly on short timescales, were also prioritised. In some cases additional targets of interest with compatible coordinates, but somewhat higher $\log(R'_{\rm HKmax})$ were then added to the corresponding observing run, to fully utilize telescope time. The basic $\log(R'_{\rm HKmax}) < -5.10$ criterion was always adhered to, so the additional prioritisations did not affect our fundamental approach of a highly targeted search around stars with anomalously low activity.

2.4.1 Absorption in the interstellar medium

As shown by Figure 2.7, the MS sample predominantly lies within ~ 80 pc of the Solar System, i.e. within the local cavity of the interstellar medium (ISM) with rarefied Ca II abundances (Welsh et al. 2010). Fossati et al. (2017) showed that in general, ISM absorption can depress $\log(R'_{\rm HK})$ by 0.05 - 0.1 dex for stars at distances ~ 100 - 500 pc, with average column densities of $\log(N_{\rm CaII}\ [{\rm cm}^{-2}]) = 12$. This is at a level exceeding typical activity measurement and calibration uncertainties. For the more distant known transiting planet hosts (Fig. 2.7), ISM absorption is therefore an important consideration and is discussed in Section 3.4.1. All 39 MS sample targets I identified lie within ~ 100 pc. At these distances, measurements along 266 lines of sight in Welsh et al. (2010) consistently show $\log(N_{\rm CaII}\ [{\rm cm}^{-2}]) < 12$, with an average of 10.6. The targets are not dramatically biased towards larger distances compared to the overall MS sample (Figure 2.9). However, it is conceivable that our lines of sight towards some or even all targets pass through rare, particularly overdense regions of the local ISM. Using the online tool developed in Fossati et al. (2017)[1] I calculated

[1] http://geco.oeaw.ac.at/software.html

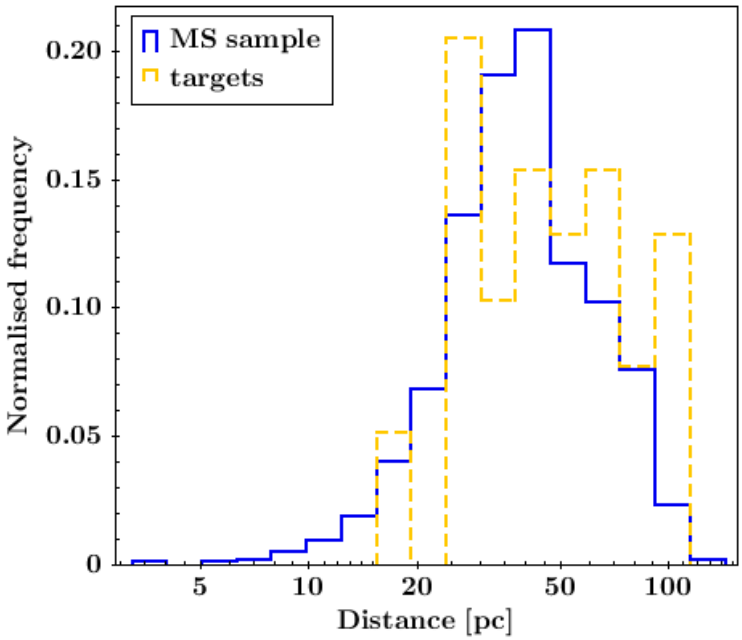

Figure 2.9: Distance distribution of my MS sample compared to the targets of interest below the basal limit.

the expected $\log(R'_{\rm HK})$ depression in "worst case scenarios": stars with intrinsic emission at exactly the basal level, and ISM absorption at velocities centred exactly on the Ca II H & K line cores (see Section 3.4.1 and Fossati et al. 2017 for details). This provides lower limits for the ISM column densities required to explain the outliers seen in Figure 2.8. Over the $B-V$ range of our targets of interest, these calculations showed that $\log(N_{\rm CaII}\ [{\rm cm}^{-2}]) \gtrsim 12.5$ - 14 is required to explain outliers at the 0.1 dex level below the basal limit.

The only direct way to assess if such anomalous ISM densities are the cause of the observed $\log(R'_{\rm HK})$ outliers is mapping of the ISM column density close to the lines of sight to the targets. This approach requires high resolution spectroscopy of early-type stars with suitable coordinates and distances (as for WASP-12 in Fossati et al. 2013). It is expensive in terms of telescope-time, and may provide no conclusive results for our targets. Early type stars at the ideal distances of <100 pc are rare. Any putative anomalous local ISM overdensities could be confined to angular sizes smaller than the separations between our

targets and suitable early-type stars. Therefore we did not carry out observations along these lines.

2.5 Young cluster targets

In addition to targets below the basal limit, circumstellar absorption may also be present around "above-basal" stars with significantly lower $\log(R'_{\rm HK})$ than expected for their age. In practise this is a difficult comparison to make, as stellar ages are unknown or at best poorly constrained for the vast majority of field stars. However several young, nearby clusters have both well established ages and $\log(R'_{\rm HK})$ measurements for dozens of cluster members. Since stars in a particular cluster form and evolve in unison, their activity is highly correlated (Mamajek & Hillenbrand 2008, hereafter MH08) allowing a search for anomalously low values within the cluster's activity distribution. An important caveat is that stellar rotation and activity have not necessarily converged from disparate initial rotation rates at formation, depending on the age of the cluster. Early-type stars reach the so-called "slow-rotator", i.e. convergent rotation sequence first, followed by late-type stars. This leads to rotation period-stellar mass distributions as in Figure 2.10. Due to this effect and the significant intrinsic activity variability of young stars, cluster activity distributions can show significant scatter. This is compounded by "interloper" measurements of the (lower) activity levels of field stars erroneously classified as cluster members.

Notwithstanding these difficulties, a search for short-period planets with high mass-loss rates in young clusters is highly interesting. As mentioned in Section 1.2, the formation and orbital evolution of short-period planets are poorly understood. Competing theoretical models in these areas are still in very early stages of observational verification. A major limitation is the small num-

ber of well-characterised exoplanets discovered around young stars and stars with well-constrained ages. Clusters members provide many benefits: their precise ages and stellar parameters and their chemical homogeneity are well-suited for studies of planet-star parameters and dependencies (e.g. Cochran et al. 2002).

Planetary migration mechanisms (Adams & Laughlin 2006) and mass-loss processes (Lammer et al. 2003) are thought to be most important in the first few hundred Myr of a system's life, so young cluster hosts ($\lesssim 1$ Gyr) of short-period planets are particularly valuable discoveries (e.g. Mann et al. 2017). There have been several efforts to discover planets in young clusters and associations through transit searches (e.g. Mochejska et al. 2002, Aigrain et al. 2007, van Saders & Gaudi 2011) and RV programs (e.g. Cochran et al. 2002, Crockett et al. 2012, Bailey et al. 2017), with relatively few discoveries of short-period planets to date. High activity levels of young stars result in elevated RV and photometric jitter (e.g. Crockett et al. 2012), making planet searches more challenging. See Section 4.2 for more details on stellar jitter in the RV planet search context. The high-precision space-based *K2* transit survey (Howell et al. 2014) recently targeted several young clusters, resulting in the discovery of planets with radii between Earth and Neptune (Mann et al. 2017).

I focussed on searching for anomalous activity values for 2 particularly well-studied clusters with abundant archival activity data: the Hyades and Pleiades. Note that at ages of $\sim$ 130 Myr (Pleiades) and $\sim$ 630 Myr (Hyades), these clusters are too old to retain primordial gas from a protoplanetary disk (c.f. Wyatt et al. 2015, Gorti et al. 2016 and references therein). Significant circumstellar absorption is not expected in the absence of a source replenishing circumstellar material. For the Pleiades, the observations in Ritchey et al. (2006) showed relatively small variability in the interstellar Ca II column density (at the $\lesssim 1$ dex level). The Hyades cluster is very nearby, with a mean distance of 46 pc

(Perryman et al. 1998). I confirmed that the interstellar Ca II column densities in the Hyades direction (out to 80 pc), as measured by Welsh et al. (2010), are low ($\log(N_{\mathrm{CaII}}\ [\mathrm{cm}^{-2}]) < 10.5$). Thefore ISM absorption should affect all cluster members to a similar, negligible degree (c.f. Fossati et al. 2017).

Like the outliers in Section 2.4, Hyades & Pleiades stars with unusually low $\log(R'_{HK})$ may therefore be hosts of short-period mass-losing planets; the only difference is the higher intrinsic activity of cluster stars. However very strong evidence of cluster membership is required, as their "above-basal" $\log(R'_{HK})$ values are only anomalous in the cluster context.

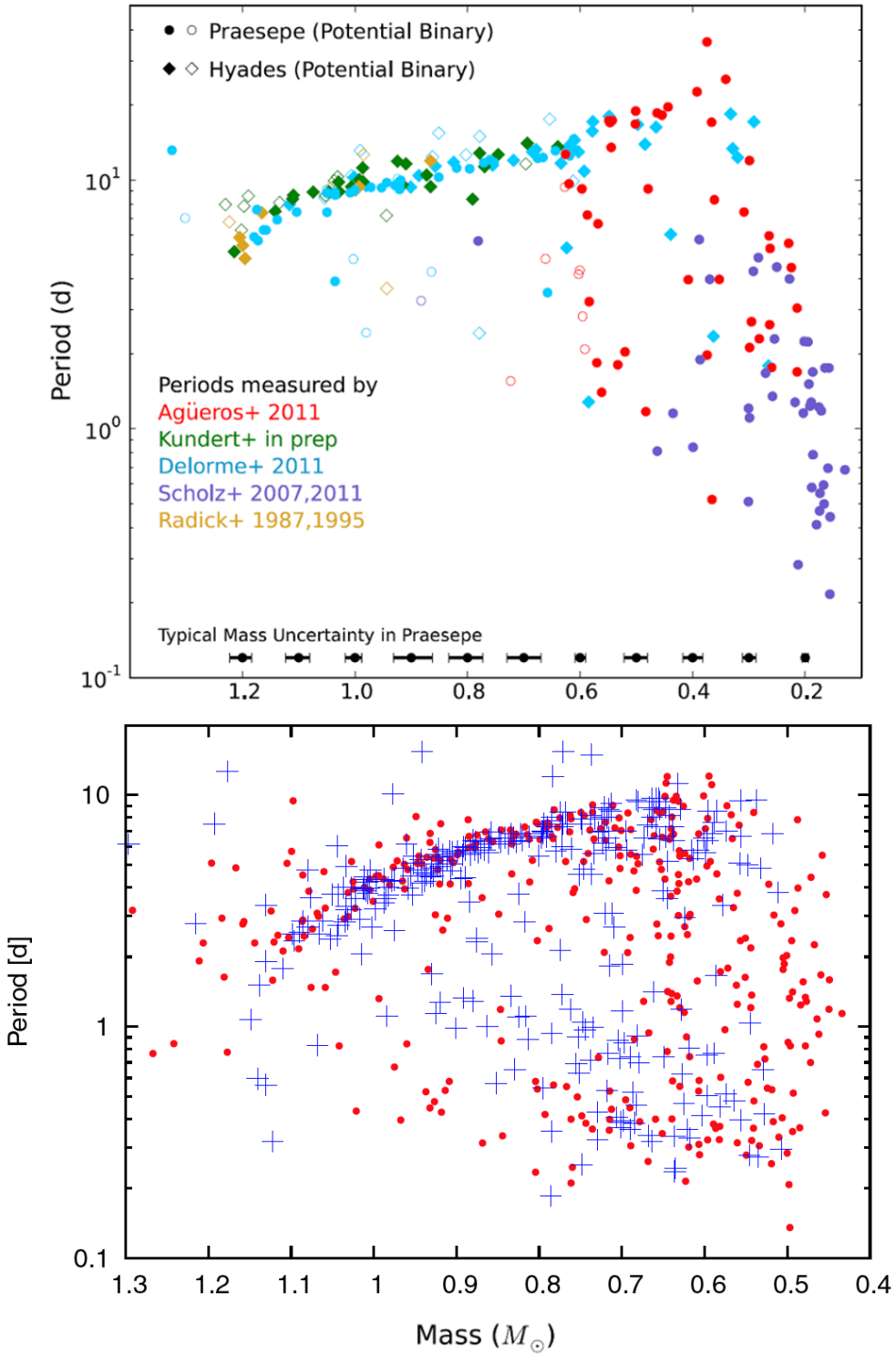

Figure 2.10: Figures taken from Hartman et al. (2010) and Douglas et al. (2014) to illustrate the stellar rotation period distributions of the Hyades (**top panel**) and the younger Pleiades and M35 clusters (dots and crosses respectively, **bottom panel**). The "slow-rotator" sequences and the ultra-fast rotators are apparent. At the Hyades age, the vast majority of early-type stars (that are not spun up via binarity) have convergent rotation periods. Stars above the sequences are likely interlopers.

2.5.1 Hyades members

The Hyades cluster is one of the best-studied stellar associations, particularly in terms of stellar activity (MH08). An initial Hyades precision RV planet search program outlined in Cochran et al. (2002) did not lead to planet discoveries, but provided detailed stellar properties (Paulson et al. 2003), activity data (Paulson et al. 2002), RV variability levels and detection limits (Paulson et al. 2004), which primarily covered the Hot Jupiter parameter space. Quinn et al. (2014) discovered the first Hyades planet (a HJ) through a new RV survey; followed by the detection of a short-period Neptune around a Hyades M-dwarf via *K2* data (Mann et al. 2016, David et al. 2016). With additional Hyades *K2* observations scheduled, further discoveries are likely (Mann et al. 2017). I note that for the 2 short-period Hyades planets discovered to date, no calibrated $\log(R'_{\rm HK})$ values have been published.

Due to the complications described in the previous section, I only used the thoroughly vetted, "well-behaved" and homogeneous dataset of multi-epoch HIRES measurements presented by Paulson et al. (2002). Their sample is pre-selected to have $v\sin i < 15$ km s^{-1}. This avoids scatter in the cluster's activity distribution arising from very fast rotators that have not yet reached the convergent rotation sequence, and ensures RV follow-up is feasible. Note that the P13 and MH08 datasets contain additional measurements of probable Hyades members. However these are from heterogeneous sources and I found several "dropout" issues as described in Section 2.2. Therefore I only used the Paulson et al. (2002) sample (hereafter P02), containing 82 Hyades stars that have a minimum of 6 HIRES activity and radial velocity measurements per star. Note that P02 published $\log(R'_{\rm HKmean})$ values, rather than individual measurements. Four of the P02 stars are flagged as likely non-members by Paulson et al. (2002) and/or MH08 based on their anomalously low activity and photometric, velocity and distance criteria (Figure 2.11). I retained one of these

stars (HD35768) as a target of interest, as its mean activity level is below the MS basal limit and it fulfills the unevolved MS criteria described in Section 2.3. Note this star is not included in the compilation of Pace (2013), so it was not flagged up in my original MS sample. Since the other 3 probable non-members do not fulfill my MS and/or $\log(R'_{HK})$ criteria, I discarded them. The P02 activity distribution is shown in Figure 2.11, using the $\log(R'_{HKmean})$ values reported in Paulson et al. (2002).

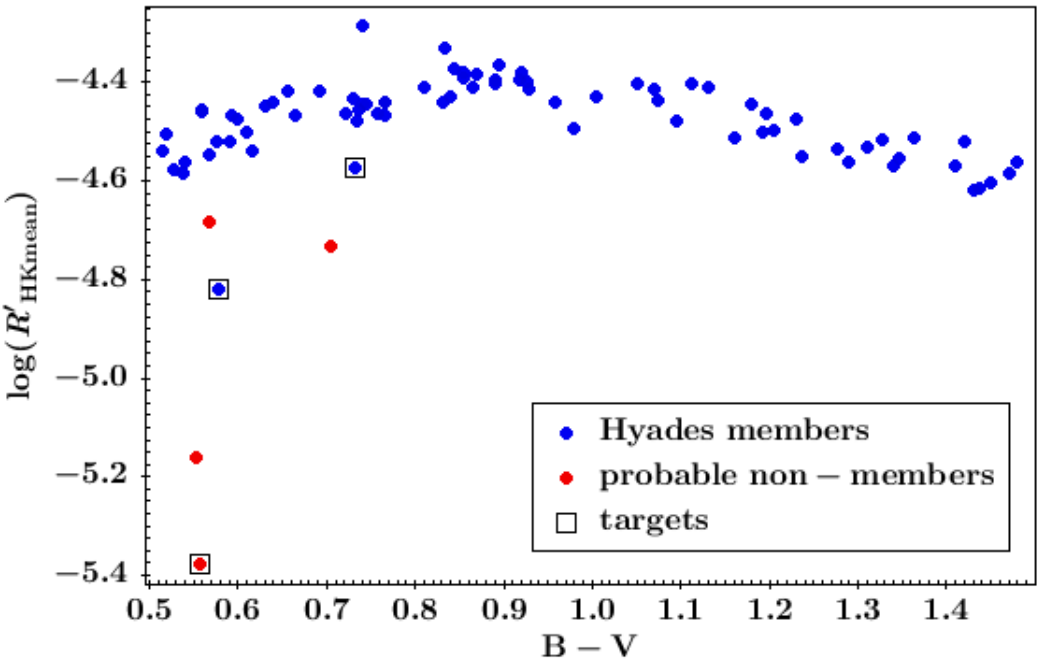

Figure 2.11: Activity values of the P02 Hyades sample against colour, highlighting probable non-members and our targets.

The rigorous membership vetting applied to the P02 sample and the use of multi-epoch averages of high-quality activity values from a single instrument result in a remarkably well defined trend of $\log(R'_{HK})$ with spectral type. As shown in Figure 2.11, after excluding probable non-members, I identified 2 possible Hyades targets below this cluster trend, HD19902 & HD30589. I verified that both are in fact convincing Hyades members. Their membership is confirmed by HIPPARCOS kinematics (Perryman et al. 1998, de Bruijne et al. 2001) and precision abundance analysis of ~ 20 elements (Paulson et al. 2003;

De Silva et al. 2006, Tabernero et al. 2012). Crucially, their Lithium abundances fall on the tight, Hyades-specific trend (Balachandran 1990;Takeda et al. 2013).

To assess these anomalies further, we re-reduced archival HIRES observations of the targets along with comparison Hyades members from P02 with similar stellar properties and activity values on the cluster trend. The comparison (Figure 2.12) confirmed that the Ca II H & K line cores of the targets are clearly depressed.

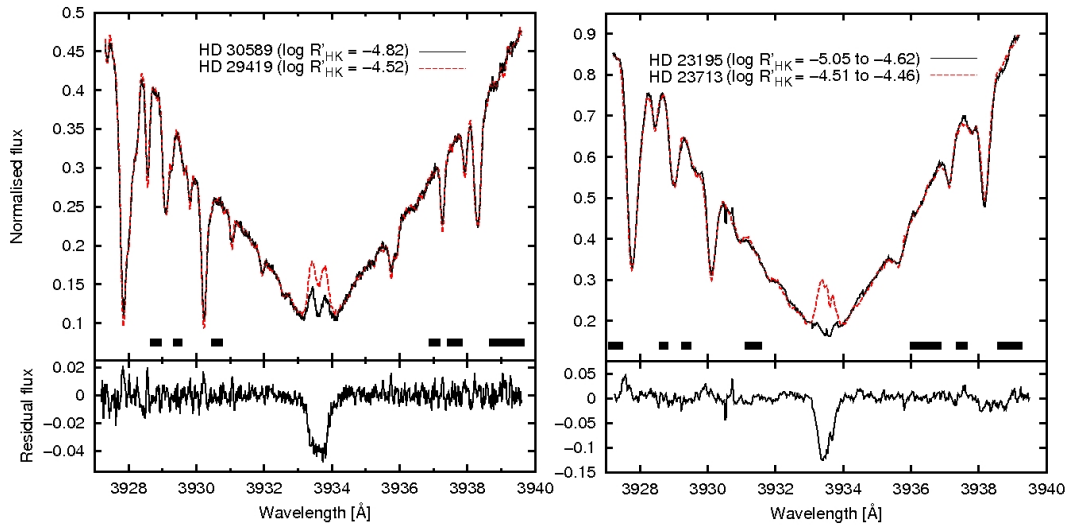

Figure 2.12: Ca II K line profiles of a Hyades (left) and Pleiades (right) target and "normal activity" comparison stars (red) of the same spectral type and from the same cluster. Spectra are normalised to the windows highlighted in the plot. **Lower panels:** target flux - comparison flux. Line profiles of the other targets in each cluster appear very similar to the examples shown.

Qualitatively, this is consistent with absorption features from low-velocity circumstellar material, but may also arise from a lower intrinsic chromospheric emission. The latter explanation would require both targets to show an anomalously low intrinsic activity level with respect to the cluster trend over the course of 10 observations covering ~4 years. To rigorously address this hypothesis we would need a comparison of their $\log(R'_{HK})$ values to other activity indicators that are thought not to be affected by circumstellar absorption. Unfortunately, such measurements have not been published for HD19902 & HD30589. Their respective $v \sin i$ values are 1.5 and 5.3 km s^{-1} (Paulson et al.

2003), consistent with the Hyades distribution of $v \sin i$ against spectral type (see Fig. 3 of Paulson et al. 2003). This disfavors a scenario where the anomalously low activity measurements arise from a pole-on view along our line of sight to these stars, foreshortening the most active lower latitude regions.

Note that for HD 30589, HD 19902 & HD 35768, a small number of radial velocity measurements were published by (Paulson et al. 2004). Their sampling is sparse, with baselines of ~ 6 years and ~ 11 observations per star. The RV RMS ranges from 6 - 16 $\mathrm{m\,s^{-1}}$, i.e. at a typical level for their entire Hyades sample. These datasets are of little use in the search for short period signals, apart from ruling out Hot Jupiter companions. A high cadence RV search over a baseline of a few nights will cover approximately a single stellar rotation cycle (see Figure 2.10). When searching for ultra-short period planets, this strategy should be significantly less sensitive to RV jitter on the timescales of stellar rotation compared to the observations of Paulson et al. (2004).

2.5.2 Pleiades members

Given its proximity (~ 130 pc, e.g. Soderblom et al. 2005) the Pleiades cluster has long been a benchmark for studies of stellar evolution, including the landmark study of stellar spin-down by Skumanich (1972). No confirmed planets have been discovered in the Pleiades to date (Mann et al. 2017), in spite of the cluster being targeted by the *K2* mission. Gaidos et al. (2017) showed that this does not constitute evidence for a difference between the underlying planet population of the Pleiades and that of field stars.

Unfortunately no sizeable homogeneous activity catalogue (comparable to P02) has been published for the Pleiades. I therefore used the compilation of MH08, which applied rigorous cluster membership vetting but combines

heterogeneous activity measurements from several sources, with both single epoch entries and multi-epoch averages. The resulting activity distribution is shown in Figure 2.13. The much greater activity scatter compared to the P02 Hyades sample is likely caused by a combination of higher intrinsic variability at the younger Pleiades age, heterogeneous activity data, and a lack of $v \sin i$ filtering as done for P02.

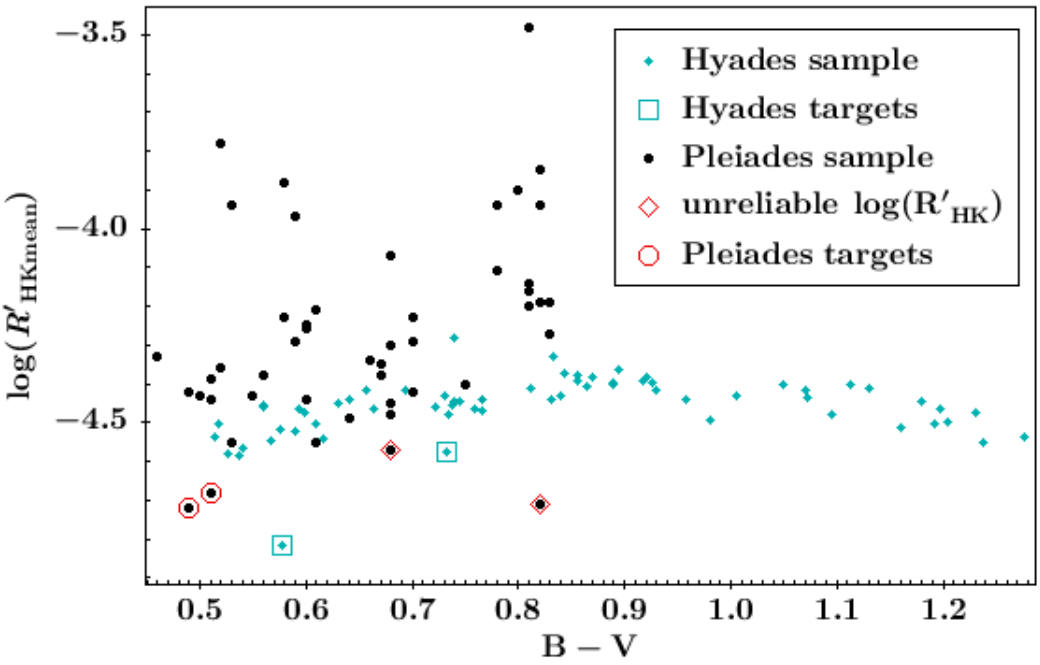

Figure 2.13: Activity values of the Pleiades sample against colour, compared to the P02 Hyades sample (excluding probable non-members). 4 potential Pleiades targets are highlighted, 2 of which were found to have potentially unreliable activity values.

As shown in Figure 2.13 I defined all Pleiades stars with activity levels below the Hyades distribution as potential targets of interest. This is a conservative criterion: their activity levels are below those of a cluster that is ~500 Myr older. Two of these targets were ruled out for follow-up work:

1. HD282975 is a double-lined spectroscopic binary, with two G-type stars in a high-eccentricity (e ~ 0.6, P = 26d) orbit (Mermilliod & Rosvick 1992). The system has been searched for circumbinary planets, with a null-result presented by Konacki et al. (2009). The authors find that the

shortest stable orbital period of any planet is 251 days and rule out Jupiter mass planets on orbits up to several years. Clearly this is not a suitable system to search for short-period planets. Its activity value should be re-examined, as unreliable measurements are plausible for double-lined binaries (e.g. Beck et al. 2016).

2. Cl Melotte 22 2027 is a system with three components (Mermilliod & Rosvick 1992), composed of a visual binary at 0.2 ″ separation. One of these components is a close (e ∼ 0.3, P = 49d) spectroscopic binary (Halbwachs et al. 2003). Again, the activity value may be unreliable and a RV search for short-period planets is not advisable in any case.

This leaves two targets of interest, HD23195 and BD+22553. Their Pleiades membership is confirmed by proper motions and (IR-) photometry (Schilbach et al. 1995; Stauffer et al. 2007). Their Lithium & Berillium abundances are undepleted, i.e. at the same level as for all Pleiades members (Boesgaard et al. 2003). This is conclusive, since destruction of these elements with stellar age occurs very rapidly at their spectral type, in the so-called "Lithium dip". We verified that the anomalously low $\log(R'_{HK})$ values were reliable, by examining archival HIRES spectra. Our comparison of the targets' Ca II line profiles to stars with "normal" Pleiades activity levels confirmed the anomalies (Fig. 2.12), similarly to the Hyades case.

Note there are no public RV timeseries for these Pleiades targets. Neither our young cluster targets nor our MS sample outliers have been observed by the *K2* mission.

2.6 Summary

I used a large compilation of archival activity values to identify objects showing both transient and consistently low outliers. Possible causes of short-lived activity outliers ("dropouts") were investigated, including planetary and cometary transit signals and spurious systematic effects. I found the latter explanation is likely for the majority of outliers, in particular those within the Isaacson & Fischer (2010) dataset. I confirmed that the only Ca II planetary transit detection reported to date in the literature has a clear effect on the corresponding activity timeseries. The ~1% activity depression in-transit is well below the 30 - 90% amplitude of the activity dropouts I identified among my large stellar sample. I then selected main sequence stars consistently below either the basal limit or the relevant cluster distribution. These are the targets for testing the hypothesis that anomalously low activity levels can provide a short-cut to identifying mass-losing, short-period exoplanets.

Details on collaborative contributions

The idea of searching for stars with anomalously low activity levels as targets for a short-period planet search originated from my supervisor C.A. Haswell. J.R. Barnes and G. Anglada-Escude contributed important analysis of the public HARPS, low SNR observations (covering known HJ transits, see Section 2.2). This led to the conclusion that the data was unsuitable for the detection of Ca II transit signals. I identified the apparent, high SNR Ca II transit signal in the HD189733 dataset. All detailed analysis of this signal and its origin (extending to other spectral lines) was then carried out by J. R. Barnes, leading to the publication of Barnes et al. (2016). I provided feedback for this manuscript. J.R. Barnes reduced and analysed the archival HIRES spectra used as a comparison for cluster stars in Section 2.5 and provided Figure 2.12.

Chapter 3

SALT observations of the Chromospheric Activity of Transiting Planet Hosts: Mass Loss and Star Planet Interactions

A significantly shortened version of this chapter was published as Staab et al. (2017).

3.1 Introduction

As detailed in Section 1.4, there has been significant interest in the chromospheric Ca II H & K line flux in HJ host stars, with the main hypotheses that:

1. A HJ planet can stimulate stellar activity through magnetic and/or tidal star planet interactions (SPI; Cuntz et al. 2000).

2. A HJ planet can suppress stellar activity through tidal interactions (Miller et al. 2012; Pillitteri et al. 2014).

3. Mass loss from a HJ planet can form a diffuse circumstellar gas cloud which absorbs in the cores of strong resonance lines (e.g. Ca II H & K and Mg II h & k) suppressing the measured stellar activity below its true value (Haswell et al. 2012; Fossati et al. 2013).

As discussed in Chapter 1, it is important to identify individual systems hosting known short period planets where mass loss appears to be masking the intrinsic activity, and cases where the stellar activity appears boosted by SPI. Ca II H & K emission enhancements from SPI should be a good predictor of a system's radio brightness (e.g. See et al. 2015). With enhancements in radio astronomy capabilities (ALMA, ALMA Partnership et al. 2015; SKA, Carilli & Rawlings 2004), detections of exoplanetary radio emissions are imminent (Vidotto et al. 2015). These are expected to yield planet rotation periods and magnetic moments, with important implications for exoplanetary magnetospheric physics and the transfer of energy and angular momentum between the host star and planet, driving orbital evolution. SPI processes themselves are active areas of research (see Section 1.1.1). Identifying additional systems with evidence for SPI can inform detailed follow-up observations for characterising SPI. Systems where the Ca II H & K emission is absorbed by gas lost from the planet offer the potential to determine the planet's chemical composition through transmission spectroscopy.

3.1.1 Importance of stellar activity data from SALT

There are few southern hemisphere spectrographs calibrated to produce $\log(R'_{\rm HK})$, and large telescopes or significant exposure times are required to achieve sufficient signal to noise in the Ca II H & K cores of typical HJ host stars (V= 11 or fainter).

At the time of writing, approximately 200 out of 2900 confirmed planet hosts (as listed on exoplanet.org) have published S-values, i.e. less than 10%. While this database is not exhaustive for activity data, it is certainly the case that the majority of exoplanet hosts lack $\log(R'_{\rm HK})$ data. Currently ~40% of the ~ 170 known transiting HJ systems (defined as $M_{\rm p} > 0.5$ M$_{\rm J}$, $P < 10$d) have known activity values, as listed by Figueira et al. (2014). The vast majority of HJs are discovered through ground-based transit surveys. WASP-S and HAT-S are responsible for almost all such discoveries in the southern hemisphere, and both routinely publish without measuring $\log(R'_{\rm HK})$ values and lack follow-up to achieve this goal. Note that only ~17 out of 75 known southern transiting HJs currently have a $\log(R'_{\rm HK})$ value published in Figueira et al. (2014).

In this chapter I report observations using the Robert Stobie Spectrograph (RSS; Kobulnicky et al. 2003) at the Southern African Large Telescope (SALT; Buckley et al. 2006), calibrating it to measure $\log(R'_{\rm HK})$, and "pilot project" results for four HJ host stars. With the calibration, the RSS can now be used as an efficient facility for stellar activity measurements: it is attached to a 10 m class telescope, enabling high SNR observations of HJ hosts within modest time allocations. Following the increased instrumental throughput in 2015, we found that ~ 400s exposures ensured SNR > 15 in the Ca II H & K cores of stars with V $\lesssim$ 13, even in poor observing conditions. Using our RSS calibration, homogeneous $\log(R'_{\rm HK})$ measurements for the majority of southern HJ hosts will be published after additional semesters of ongoing SALT observations.

This is particularly useful because previously published data is heteroge-

neous, measured with a variety of spectrographs (see Chapter 2) and often lacks rigorous uncertainty estimates. For high SNR observations, the different instruments' calibrations to the Mount Wilson system can dominate the uncertainty on $\log(R'_{\rm HK})$ (see Jenkins et al. 2011 and Section 3.3.3). Analysis of a large, consistent RSS dataset, using the methodology described in this chapter, will not suffer these disadvantages.

In this chapter, Section 3.2 describes observations and data reduction; Section 3.3 describes the RSS calibration; Section 3.4 discusses the planet host measurements in the context of large stellar samples and the three hypotheses listed above; Section 3.5 gives conclusions and implications for future work.

3.2 Observations and reduction

All observations were taken with the RSS, a multimode instrument at SALT. It employs a set of slit masks and volume phase holographic transmission gratings and an articulating camera/detector system, allowing a choice of wavelength coverage and spectral resolution. Jenkins et al. (2011) assessed the effect of spectral resolution on the precision of $\log(R'_{\rm HK})$ measurements, finding a resolving power of R > 2500 is needed. We selected RSS long-slit spectroscopy settings to achieve the highest possible resolution in the region of the Ca II H & K lines: a 0.6''slit, and the PG3000 grating. We used the RSS simulator tool[1] to ensure the wavelength range of interest was away from the gaps between the 3 CCD detectors, choosing a camera station angle of 79.75° and grating angle of 39.875°. This yielded coverage from 3882 Å to 4614 Å, and a resolution R ~ 7300 at 4000 Å. Prebinning of 2 in the spectral direction gave 0.23 Å per binned pixel. Future $\log(R'_{\rm HK})$ measurements with the RSS using our calibration should adopt the same setup to ensure consistency. Note the SALT telescope operates in queue-schedule mode. We prepared observation

[1] *http://astronomers.salt.ac.za/software/rss−simulator/*

blocks for each target with the instrument settings described, and the data was collected by the local SALT astronomers.

3.2.1 Target selection

To permit calibration onto the Mount Wilson system, we selected calibration stars from Baliunas et al. (1995), hereafter B95. B95 reported both the variability and the mean of S-values, $< S_{\rm MW} >$, using decades of data from the Mount Wilson survey. We included stars observable with SALT, ranging from very inactive, $< S_{\rm MW} > = 0.14$, to very active, $< S_{\rm MW} > = 0.5$, (Table 3.2), chosen for low variability and $0.4 < B-V < 1.0$. Note HD 10700 and HD 22049 showed stable activity for a decade after B95's observations (Hall et al. 2007). The calibrator stars are bright, so we took 4 - 6 consecutive, short (15s - 70s), high SNR exposures, to avoid saturation.

Our four science targets, WASP-43, WASP-51/HAT-P-30, WASP-72, and WASP-103 (Hellier et al. 2011; Johnson et al. 2011; Gillon et al. 2013; Gillon et al. 2014; Table 3.1), are host stars of transiting HJ planets. They were all discovered after the publication of Knutson et al. (2010) and are notable for proximity to their host stars. WASP-103 b suffers appreciable tidal deformation due to proximity to the host star, and almost fills its Roche lobe (Figure 3.1). In this respect, WASP-103b is similar to WASP-12b (c.f. Figure 21 of Haswell et al. 2012) and we chose to make multiple exposures of this target to search for possible variations in circumstellar absorption of Ca II H & K with orbital phase. To help with the already highly constrained SALT scheduling, we did not specify particular observing phases, but allowed our WASP-103 visits to fall as determined by the queue-scheduling.

Tables 3.2 and 3.3 summarise the observations, omitting three WASP-103 spectra which suffered cosmic ray hits directly in the Ca II H & K lines. We discarded these, as any activity measurement would be unreliable: low-level

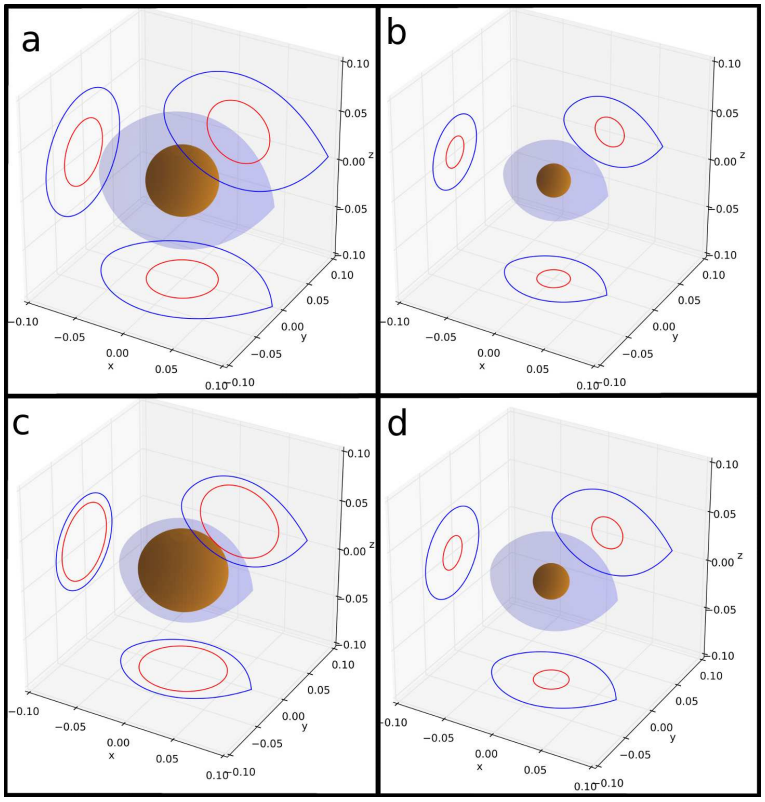

Figure 3.1: The Roche lobes (blue) of the HJs (orange) WASP-43, WASP-51, WASP-103 & WASP-72 (a-d), calculated using stellar and planetary parameters from TEPCAT (Southworth 2011). We corrected the empirical radius for the projection effect of a non-spherical planet viewed at orientation $i \neq 90°$, to give a self-consistent 3-D Roche geometry. Axes are in units of orbital separation, with the x and y axis in the orbital plane and z perpendicular to it. For the projections, red outlines represent the planet and blue outlines the Roche lobe.

artefacts or residuals left by cosmic ray cleaning in the few pixels of the core bandpasses may significantly change the extracted $\log(R'_{\rm HK})$ value. Exposure times ranged from 320s to 2030s for the planet hosts, ensuring SNR > 15 in the Ca II H & K cores of the extracted spectra. The calibrators HD 26913 and HD 26923 are separated by ~ 40″, and were therefore observed simultaneously by appropriately orientating the RSS slit.

Table 3.1: Stellar and planetary parameters relevant for this work, taken from TEPCAT (Southworth 2011), including stellar effective temperature ($T_{\rm eff}$), semi-major axis (a), orbital period, planetary equilibrium temperature ($T_{\rm eq}$) and orbital inclination i.

System	$T_{\rm eff}$ [K]	a [AU]	Period [d]	$M_{\rm p}$ [$M_{\rm J}$]	$R_{\rm p}$ [$R_{\rm J}$]	$g_{\rm p}$ [m s^{-2}]	$T_{\rm eq}$ [K]	i [deg]
HAT-P-30/ WASP-51	6338 ± 42	0.042	2.81	0.71 ± 0.03	1.34 ± 0.07	9.8 ± 0.9	1630 ± 42	83.6
WASP-43	4520 ± 120	0.015	0.81	2.03 ± 0.05	1.04 ± 0.02	47.0 ± 1.4	1440 ± 40	82.3
WASP-72	6250 ± 100	0.037	2.23	1.46 ± 0.06	1.27 ± 0.20	22.9 ± 7.3	2210 ± 120	81.8
WASP-103	6110 ± 160	0.020	0.93	1.47 ± 0.11	1.55 ± 0.05	15.1 ± 0.9	2495 ± 66	88.2

3.2.2 Data reduction

The SALT pipeline (Crawford et al. 2010) corrected all our science and calibration exposures for CCD bias, gain and crosstalk between the CCD amplifiers. We then performed flat fielding, cosmic ray cleaning and wavelength calibration in IRAF, using arc-lamp exposures taken immediately after each target. This minimises variability in the wavelength scale, caused by spectrograph flexures, and is the default calibration strategy for the RSS[2]. The Copper-Argon (CuAr) arc-lamp was selected to provide sufficient calibration lines at the critical blue wavelength end near Ca II H & K. Spectra were extracted with particular attention to reliable background subtraction, since over- or undersubtraction biases $\log(R'_{\rm HK})$ measurements (Fossati et al. 2017). I describe individual reduction steps in more detail below.

Flat-fielding and cosmic ray cleaning

We median-combined the set of flat-fields taken closest in time to each of the science observations. Nightly flats were not consistently provided by the SALT team, as their experience showed the flatfield is sufficiently stable on timescales of weeks. Note that a column of "lazy pixels" on the detector at $\sim$ 3994 Å falls into the R continuum passband needed for $\log(R'_{\rm HK})$ measurements (Sect. 3.3). Flatfielding was crucial to correct for this artefact and ensure reliable activity measurements. For the longer exposures in the dataset, a significant number of artefacts from cosmic rays hits (CR) are present. CRs within the stellar spectrum and the regions used for background fitting are accounted for during the final spectral extraction step. However, we found that a smoother overall reduction process was achieved by adding an initial CR cleaning step (after flat-fielding). We used the *L.A.Cosmic* algorithm (van Dokkum, 2001, PASP, 113, 1420) for this.

[2] *http : //pysalt.salt.ac.za/proposalcalls/current/ProposalCall/*

Wavelength calibration

Exposures taken for longslit spectroscopy generally suffer from geometric distortions. This is illustrated by the arc-lamp exposure shown in Fig. 3.2: the CuAr emission lines have substantial curvature relative to the CCD columns. A two-dimensional wavelength solution is therefore needed to calibrate an entire science frame. We used the IRAF package *twodspec.longslit* for this process:

1. A one-dimensional wavelength solution was fitted at a single position along the spatial axis, using the IRAF *identify* task and the CuAr line-list provided by the SALT team[3]. We found that 3rd order cubic splines provided solutions with both low RMS (typically $\lesssim 0.03$ Å) and no systematic structure in the residuals.

2. The IRAF task *reidentify* then re-fitted this wavelength solution at 10 pix increments along the spatial axis of the CuAr frame, tracing out the curvature of each emission line.

3. The output from the previous steps was used by *fitcoords* to compute the surface defining wavelength as a function of x & y position on the CCD frames.

4. Finally, the *transform* task and the *fitcoords* output was used to geometrically correct exposures, so that wavelength is a linear function of the x-axis, and position along the slit is linear along the new y-axis. This is best visualised by a transformed arc-lamp exposure, shown in Fig. 3.2. The transformation of science frames becomes important for reliable background subtraction during extraction. It ensures that the background fit, applied at significant spatial separation from the stellar spectrum, is subtracted from the matching wavelength range on the stellar trace. This

[3] *http : //pysalt.salt.ac.za/lineatlas/CuAr.txt*

2D wavelength solution is also vital for observations with several targets along the slit, namely HD 26913 + HD 26923 exposures.

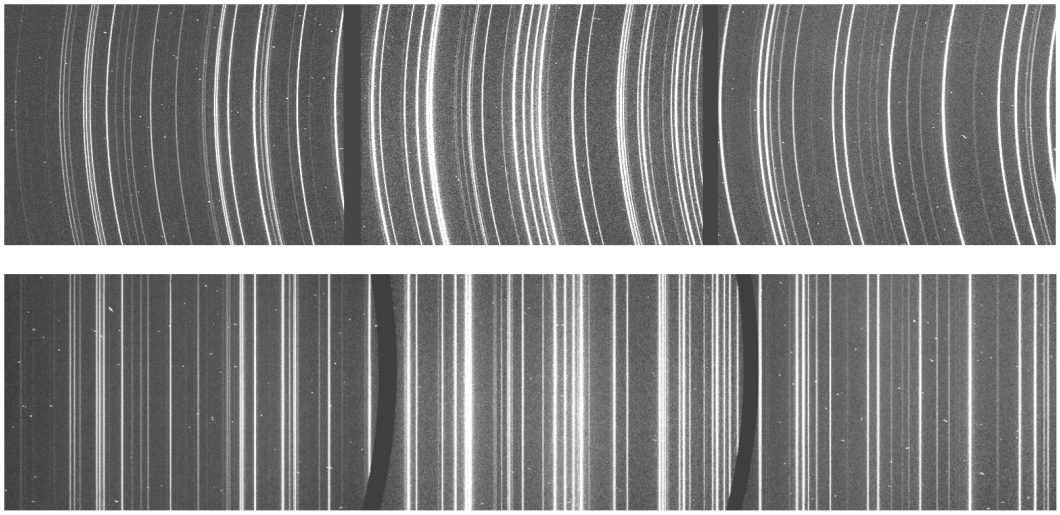

Figure 3.2: Section of an arc-lamp exposure, highlighting geometric distortions (**top**). These are removed after wavelength calibration and transformation of the arc-lamp exposure (**bottom**). The gaps between the 3 CCDs are visible.

Extraction and background subtraction

We extracted spectra using IRAF's *apall* task, using the option of "optimal", i.e. variance-weighted extraction (Horne 1986). This includes tracing of the spectra across the CCD frames and rejection of CR affected pixels in the stellar extraction and background fit windows. *apall* provided the uncertainties on the final spectrum, propagating both photon and detector readout noise per pixel. The sky background levels varied over the programme, as observations were taken in both dark and bright time and under variable cloud cover. Scattered moonlight, i.e. essentially the solar spectrum, dominated the background. Exposures of HD 182101 and a subset of WASP-103 observations contained particularly high levels of scattered light. We ensured that the background subtraction was also implemented correctly in these cases. As the background intensity distribution across the science frames also varied, quadratic and linear fits to the background were required for different exposures.

Issues of note encountered during reduction

1. Initial trimming of the frames was necessary because the arc-lamp exposures were too faint for reliable wavelength solutions at the top and bottom edges of each frame (i.e. in the spatial direction). Since these regions contained no useful areas on the science frames, all exposures could be trimmed accordingly, allowing more reliable 2D wavelength solutions for the remaining areas.

2. Interpolation across the two CCD gaps (see Fig. 3.2) was done at the start of the reduction process, as this reduced glitches in later steps, e.g. mis-identification of cosmic rays along the sharp gap boundaries.

3. Due to user-error by a local SALT astronomer the Argon arc-lamp was selected for the WASP-72 observation. This lamp has fewer emission lines than CuAr, particularly bluewards of ~ 4000 Å. Fitting the wavelength solution with the Argon line-list provided by the SALT team[4] gave a wavelength solution with satisfactory RMS of 0.016Å. I verified that the resulting WASP-72 spectrum did not suffer serious calibration problems in the wavelength region of interest for $\log(R'_{\rm HK})$, by comparing it with the RSS spectrum of HIP110785 (Fig. 3.3). This is the Mount Wilson calibrator with the most similar basic stellar properties to WASP-72, and its wavelength solution was derived via the CuAr lamp.

[4] *http : //pysalt.salt.ac.za/lineatlas/Argon_hires.txt*

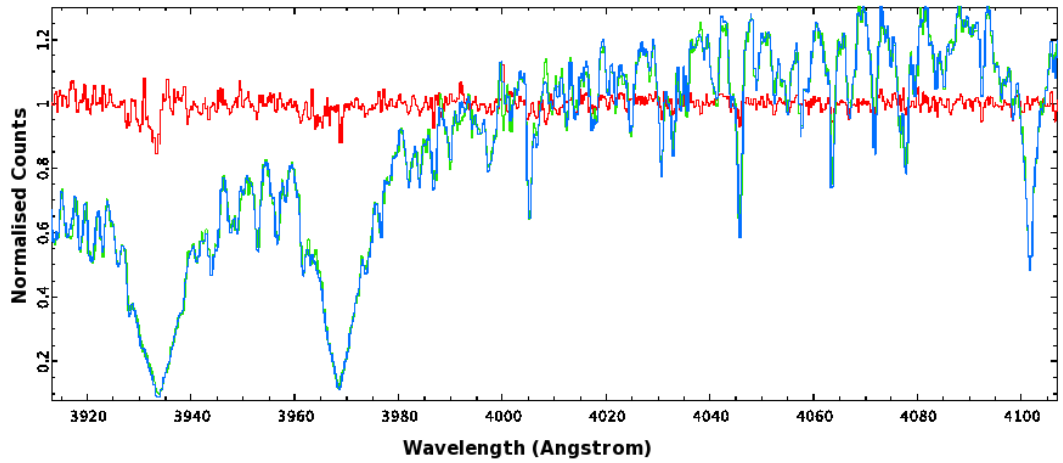

Figure 3.3: Comparison of HIP 110785 (green) and WASP-72 (blue) RSS spectra, along with their ratio (red). No obvious distortions in the WASP-72 wavelength solution are present. Note the Ca II H & K core flux depressions for WASP-72.

3.3 Analysis

3.3.1 Extraction of instrumental *S*-values

Extracted spectra were shifted into the stellar rest frame as follows. We used PyAstronomy's *crosscorrRV* routine[5] to cross-correlate each spectrum with the National Solar Observatory solar spectrum (Kurucz et al. 1984), degraded to the RSS spectral resolution. The resulting cross-correlations functions were fitted with Gaussians, as illustrated in Figure 3.4. Each spectrum's wavelength scale was then shifted by the radial velocity of the corresponding Gaussian centroid, using PyAstronomy's *dopplerShift* routine.

We then performed synthetic photometry using the Ca II H & K core bandpasses, H and K, and continuum windows, R and V, shown in Fig. 3.5. V and R are the mean flux values in 20 Å wide continuum windows centered on 3901.07 Å and 4001.07 Å while H and K are the core bandpasses, centered on 3933.664 Å and 3968.47 Å, triangularly weighted with a FWHM of 1.09 Å. We followed Lovis et al. (2011) in using the (weighted) mean bandpass fluxes, instead of integrating the flux. This reduces edge effects from finite pixel size

[5]https://github.com/sczesla/PyAstronomy

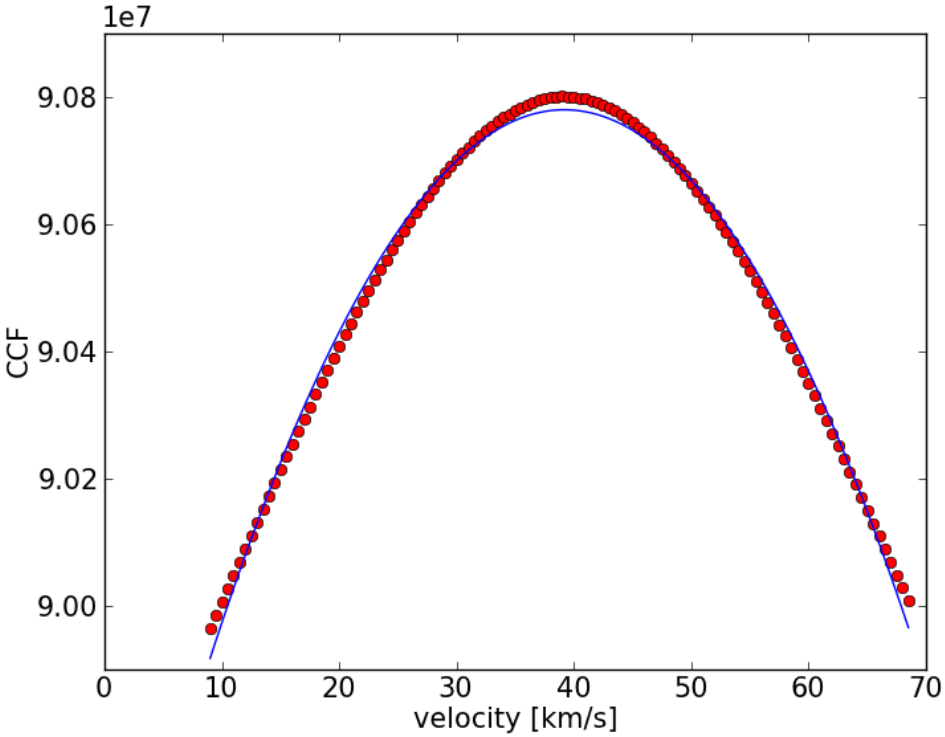

Figure 3.4: Cross-correlation function (CCF) for an RSS spectrum of HIP 12114 (red), evaluated at 0.5 km/s velocity increments, and corresponding Gaussian fit (blue).

at the band boundaries. It also renders a historical scaling factor obsolete that appears in other conversions to the Mount Wilson system. To further mitigate edge effects we rebinned all spectra to a common wavelength scale with 0.05 Å bin width, while conserving flux, using the Barak *rebin* routine[6]. The instrumental S-index, S_{RSS}, is then simply given by

$$S_{\mathrm{RSS}} = \frac{H + K}{R + V} \,. \tag{3.1}$$

Note flux-calibration is unnecessary for S-index measurements: the bandpass window placement makes these insensitive to the local spectral slope (Gray et al. 2003, Fossati et al. 2017).

[6]https://github.com/nhmc/Barak

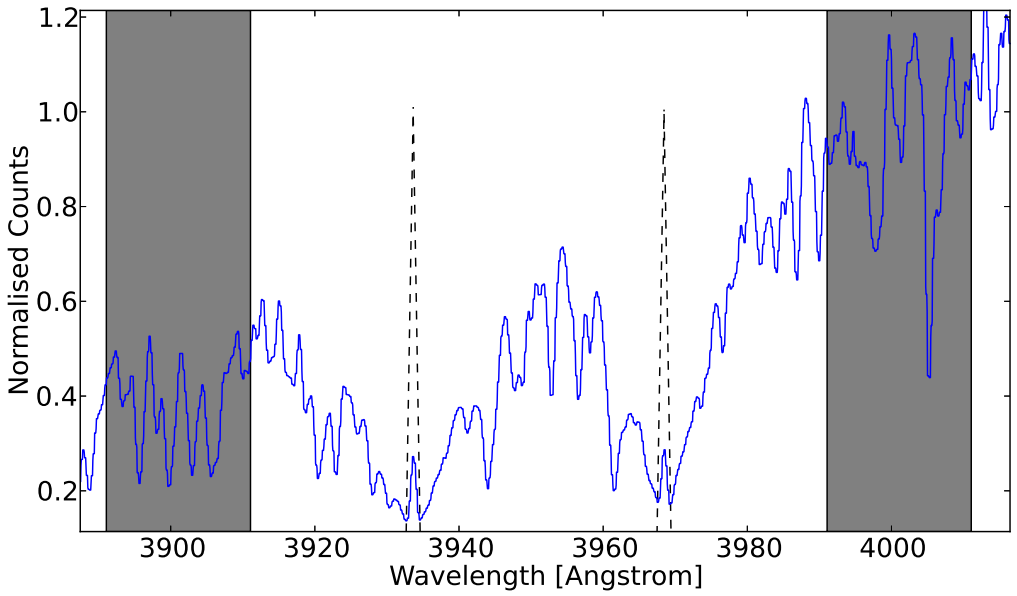

Figure 3.5: An RSS spectrum of our most active calibrator star, HD 22049, with continuum windows (greyed) and triangular core bandpasses highlighted. Strong Ca II H & K emission cores are evident. Counts are normalised to the mean of the red continuum bandpass.

3.3.2 Calibration to Mount Wilson system

Table 3.2 reports the calibration measurements with propagated photon noise uncertainties on $S_{\rm RSS}$. For the B95 values, formal measurement uncertainty on $< S_{\rm MW} >$ is negligible compared to astrophysical stellar activity variation. The current activity level of our calibrators is uncertain so I report the variability range measured over the multi-decade baseline of B95 in Table 3.2 and Figure 3.6. The relative activity variability observed by B95 typically increases with the mean stellar activity level (see also Lovis et al. 2011).

Following Jenkins et al. (2011), I performed a simple linear calibration to the Mount Wilson system, using orthogonal distance regression (ODR) via the SCIPY *ODR* algorithm[7]. ODR accounts for uncertainties in both the explanatory and response variable ($S_{\rm RSS}$ and $< S_{\rm MW} >$ in this case). I obtained (Figure 3.6):

$$S_{\rm RSS} = (0.60 \pm 0.02)\, S_{\rm MW} + (0.075 \pm 0.005). \quad (3.2)$$

[7]https://docs.scipy.org/doc/scipy/reference/odr.html

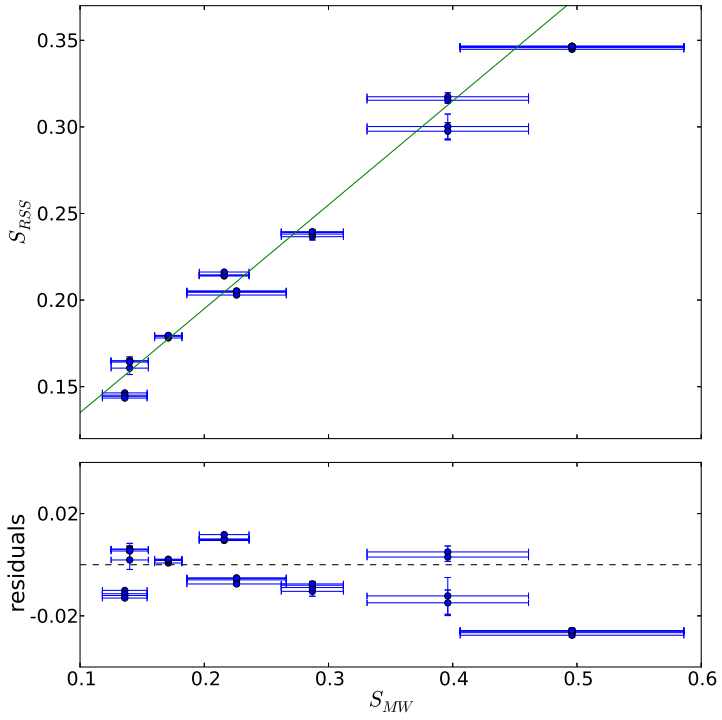

Figure 3.6: Calibration from instrumental to Mount Wilson S-values. Individual datapoints for each calibrator star overlap closely, and in most cases have S_{RSS} uncertainties smaller than the symbol sizes.

Note that the high activity calibrators carry larger uncertainties and consequently less weight in the fit. The RMS of the S_{RSS} residuals to the fit, 0.013, is at the level found in calibrations of other instruments (see Jenkins et al. 2006; Jenkins et al. 2008; Arriagada 2011). Scatter at the ~ 5 % level is a consequence of the long-term activity variation of calibration stars, combined with the single-epoch measurements of their S_{RSS} values. In principle, the scatter may be reduced as follows: instead of using the long-term average $< S_{\mathrm{MW}} >$ values from B95, the calibrators' activity cycle states at the current epoch could be extrapolated from their multi-decade B95 timeseries. The reliability of such extrapolation would have to be carefully assessed. Alternatively, the uncertainty of the calibration to Mount Wilson can be rendered negligible by investing telescope time into a large number of calibrator observations (~100), as done in e.g. Wright et al. (2004).

Table 3.2: Calibrator star measurements and associated photon noise uncertainties. Mean Mount Wilson *S*-values and their variability range were taken from B95. Note that HD 26913 and HD 26923 were observed simultaneously.

Name	$< S_{MW} >$	S_{RSS}	HJD-2450000
HD 10700	0.171 ± 0.011	0.1781 ± 0.0001	6658.339
		0.1791 ± 0.0001	6658.340
		0.1794 ± 0.0001	6658.341
		0.1796 ± 0.0001	6658.342
HD 182101	0.216 ± 0.020	0.2139 ± 0.0004	6792.586
		0.2145 ± 0.0004	6792.587
		0.2143 ± 0.0004	6792.587
		0.2162 ± 0.0004	6792.588
HD 9562	0.136 ± 0.018	0.1444 ± 0.0002	6652.321
		0.1464 ± 0.0002	6652.322
		0.1451 ± 0.0002	6652.323
		0.1434 ± 0.0002	6652.324
HD 22049	0.496 ± 0.090	0.3465 ± 0.0009	6984.514
		0.3464 ± 0.0010	6984.514
		0.3465 ± 0.0011	6984.515
		0.3448 ± 0.0009	6984.515
		0.3458 ± 0.0006	6984.515
		0.3466 ± 0.0007	6984.516
HD 26913	0.396 ± 0.065	0.3002 ± 0.0072	6985.407
		0.2975 ± 0.0050	6985.407
		0.3174 ± 0.0023	6985.408
		0.3154 ± 0.0017	6985.409
HD 26923	0.287 ± 0.025	0.2390 ± 0.0015	6985.407
		0.2366 ± 0.0019	6985.407
		0.2381 ± 0.0010	6985.408
		0.2395 ± 0.0006	6985.409
HIP 110785	0.140 ± 0.015	0.1649 ± 0.0013	6803.656
		0.1649 ± 0.0011	6803.656
		0.1642 ± 0.0030	6803.657
		0.1607 ± 0.0037	6803.658
HIP 12114	0.226 ± 0.040	0.2050 ± 0.0005	6627.328
		0.2045 ± 0.0005	6627.328
		0.2029 ± 0.0005	6627.329
		0.2053 ± 0.0005	6627.330

Table 3.3 lists S_{MW} for the planet hosts, calculated from Equations 3.1 and 3.2, and Figure 3.7 shows the Ca II H & K lines of our planet hosts. For the

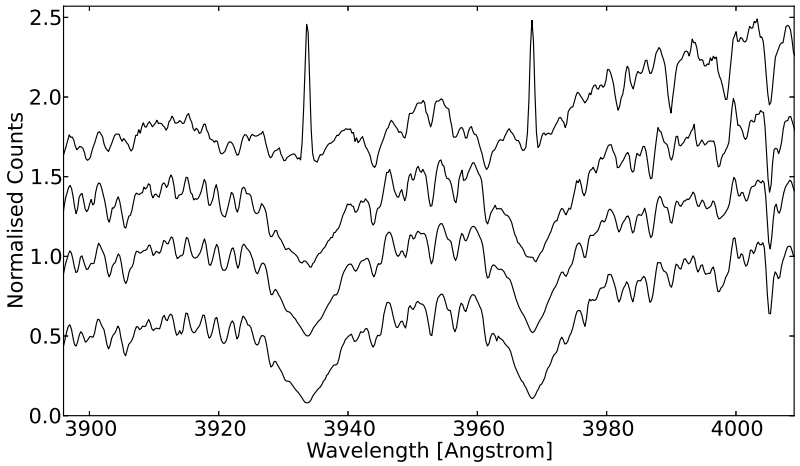

Figure 3.7: Spectra plotted in order of decreasing apparent activity from top to bottom: WASP-43, WASP-103, WASP-51, WASP-72. An arbitrary offset was added to the spectra for visibility. Counts are normalised to the mean of the red continuum bandpass.

final conversion to $\log(R'_{\mathrm{HK}})$ following Noyes et al. (1984), the stellar $B-V$ is required. No $B-V$ measurements for WASP-103 have been published; Simbad[8] and the Exoplanet Orbit Database[9] (Han et al. 2014) report $B-V$ colors for WASP-72 and WASP-51 that are completely inconsistent with detailed spectral analyses in the discovery papers (Gillon et al. 2013 and Johnson et al. 2011). For consistency, I calculated $B-V$ for all planet hosts (Table 3.3) from the stellar parameters reported in TEPCAT, using Equation 3 in Sekiguchi & Fukugita (2000) and propagating the uncertainties on the stellar T_{eff}, metallicity, and surface gravity.

[8]http://simbad.u-strasbg.fr/simbad/
[9]http://exoplanets.org/

Table 3.3: Planet host measurements and activity values derived from them. Note instrumental S-values report photon noise uncertainties, while $S_{\rm MW}$ includes the calibration uncertainty. $B-V$ values and uncertainties are derived as described in Section 3.3. Orbital phases were calculated from ephemerides in TEPCAT, with phase 0 representing mid-transit. Stellar ages in the literature are compared with activity ages derived from the MH08 relation, using our $\log(R'_{\rm HK})$ values.

Name	$\rm HJD_{-2450000}$	phase	$t_{\rm exp}$[s]	$S_{\rm RSS}$	$S_{\rm MW}$	$B-V$	$\log(R'_{\rm HK})$	activity age[a]	literature age[b]
WASP-43	6689.420	0.661	2032	1.209 ± 0.008	1.889 ± 0.065	1.10 ± 0.06	-4.17 ± 0.10	45 ± 40 Myr	300-600 Myr(1)
WASP-51	6698.410	0.883	1108	0.168 ± 0.001	0.155 ± 0.010	0.49 ± 0.01	-4.98 ± 0.07	6.2 ± 1.3 Gyr	0.5-1.8 Gyr(2)
									0.6-1.0 Gyr(3)
WASP-72	6606.307	0.332	1108	0.149 ± 0.001	0.124 ± 0.009	0.48 ± 0.03	-5.30 ± 0.15	11 ± 0.8 Gyr	2.6-3.8 Gyr(4)
WASP-103	6779.562	0.701	317	0.224 ± 0.003	0.248 ± 0.013	0.54 ± 0.05	-4.59 ± 0.04	950 ± 230 Myr[c]	1.8-6.2 Gyr(5)
	6779.566	0.706	317	0.221 ± 0.003	0.244 ± 0.013		-4.60 ± 0.04		3-5 Gyr(6)
	6779.573	0.713	317	0.221 ± 0.003	0.243 ± 0.013		-4.60 ± 0.04		
	6780.489	0.703	317	0.227 ± 0.003	0.253 ± 0.013		-4.57 ± 0.04		
	6780.493	0.707	317	0.227 ± 0.003	0.254 ± 0.013		-4.57 ± 0.04		
	6780.497	0.712	317	0.233 ± 0.003	0.263 ± 0.013		-4.55 ± 0.03		
	6780.501	0.716	317	0.230 ± 0.003	0.258 ± 0.013		-4.56 ± 0.04		
	6780.535	0.753	390	0.220 ± 0.009	0.241 ± 0.019		-4.61 ± 0.06		
	6790.473	0.490	390	0.228 ± 0.003	0.256 ± 0.013		-4.57 ± 0.04		
	6790.478	0.496	390	0.229 ± 0.003	0.257 ± 0.013		-4.56 ± 0.04		
	6790.482	0.500	390	0.235 ± 0.003	0.267 ± 0.013		-4.54 ± 0.03		
	6790.487	0.505	390	0.237 ± 0.003	0.269 ± 0.013		-4.53 ± 0.03		
	6790.502	0.521	390	0.227 ± 0.003	0.254 ± 0.013		-4.57 ± 0.04		
	6790.507	0.527	390	0.234 ± 0.003	0.264 ± 0.013		-4.54 ± 0.03		
	6790.511	0.531	390	0.225 ± 0.003	0.251 ± 0.013		-4.58 ± 0.04		
	6790.516	0.537	390	0.231 ± 0.003	0.260 ± 0.013		-4.55 ± 0.03		

[a] Uncertainty propagated from uncertainty in $\log(R'_{\rm HK})$. Note MH08 reports activity-age scatter at the 60% and 30 % levels for ages below and above 130 Myr respectively.

[b] (1)Hellier et al. 2011,(2)Johnson et al. 2011, (3)Bonfanti et al. 2015, (4)Gillon et al. 2013, (5)Southworth et al. 2014, (6)Gillon et al. 2014

[c] from mean of our measurements

3.3.3 Uncertainty budget

As shown by Table 3.2 and Figure 3.6, the $S_{\rm RSS}$ values measured for separate exposures of the same calibrator star agree to within 1%, with the single exception of 3% for HD 26913. This demonstrates that our reduction and extraction steps are robust, and indicates excellent instrumental stability, at least on short timescales. With SNR ~ 1500 at 4000 Å, there is a very small photon noise contribution to the calibration star $S_{\rm RSS}$ values. When I divided the RMS of the $S_{\rm RSS}$ values for each calibrator by the corresponding mean photon noise, a sharp rise above unity occurs below a photon noise level of ~ 0.2 % (Figure 3.8).

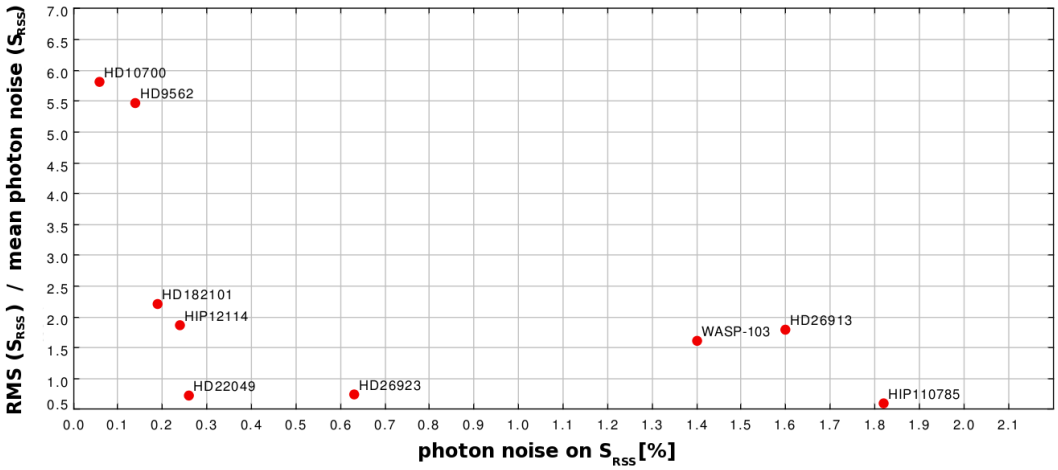

Figure 3.8: Illustration of the noise floor for our $S_{\rm RSS}$ measurements. All targets with multiple exposures are plotted. Variability above the photon noise clearly dominates the dispersion of measurements below a photon noise level of ~ 0.2 %. Due to the relatively small number targets and exposures per star, this is only an approximate estimate.

This approximate noise floor could be driven by astrophysical ultra-short timescale variability in the chromospheric emission (Narain & Ulmschneider 1996). Systematic uncertainties from instrumental stability and reduction steps are also a plausible explanation. To put our precision in context, note that Lovis et al. (2011) reached a 7-year timescale S-index precision of 0.35 % for τ Ceti with the ultra-stable HARPS spectrograph. Our $S_{\rm MW}$ (Table 3.3) have uncertainties of 3 - 8%, largely dominated by the uncertainty from the calibration relation. Clearly the noise floor contribution of ~ 0.2 % is negligible in this context. The

$\log(R'_{\rm HK})$ values listed in Table 3.3 incorporate the propagated errors on $B-V$, which is a significant contribution for WASP-43 which has a poorly constrained effective temperature. Obviously our $\log(R'_{\rm HK})$ values are snapshots during activity cycles and short-term variability. Gomes da Silva et al. (2014) used a sample of 271 mostly inactive stars, obtained over a ~ 10 year timescale, to show that astrophysical variation in $\log(R'_{\rm HK})$ results in an RMS of typically 0.015 dex, and at most 0.08 dex. Our uncertainties render the typical astrophysical variability negligible. If future, repeated observations of our 4 HJ hosts do show significant $\log(R'_{\rm HK})$ variability, this should be accounted for in the uncertainty on their mean activity levels.

3.4 Discussion

3.4.1 Activity values in context

Figure 3.9 shows our new $\log(R'_{\rm HK})$ measurements in the context of other transiting, short-period planet hosts (Figueira et al. 2014), and two samples derived from P13 data. As before, I limit all samples to $0.4 < B-V < 1.2$ (see Section 2.1.1). Due to the spurious outliers described in Section 2.2, I exclude all stars with activity values from Isaacson & Fischer (2010) from the P13 sample.

The Figueira et al. (2014) sample includes single epoch and mean activity values. I therefore used $\log(R'_{\rm HKmean})$ values derived from the P13 $S_{\rm mean}$ data (c.f. Chapter 2) for the following comparison. As described in Section 2.1.1, main sequence, subgiant and giant stars have very different activity distributions. I define an unevolved main sequence population as in Section 2.1.1, including stars < 0.45 mag above the average MS. The main sequence basal activity limit of $\log(R'_{\rm HK})$>-5.1 applies only to this sample (Wright 2004), seen in Fig. 3.9a. Highly evolved stars follow a completely different colour-activity distribution than dwarfs and subgiants. Since none of the planet hosts are highly evolved,

I reject objects more than 2.0 mag above the MS, and define a sample of moderately evolved stars between 0.45 and 2.0 mag above the MS, comprised mostly of subgiants (Fig. 3.9b). These have lower activity values than their unevolved counterparts. For most planet hosts, precise parallaxes are still unavailable, so I cannot distinguish between MS and moderately evolved stars in the same way. I therefore compare planet hosts to both populations in Figure 3.9. Note this should be revisited once precise GAIA parallaxes become available for the majority of hosts (likely in the second GAIA release; late 2017). I checked that both the MS and the evolved samples do not contain known short period planets, by cross-matching all entries with the Exoplanet Orbit Database (Han et al. 2014). There are a small number of planet hosts (15 and 19 respectively) in these samples, and the shortest orbital period is 26 days. However, the presence of planets with shorter periods cannot be excluded for the vast majority of stars in these samples, which have not been targeted with high-cadence, precision RV searches.

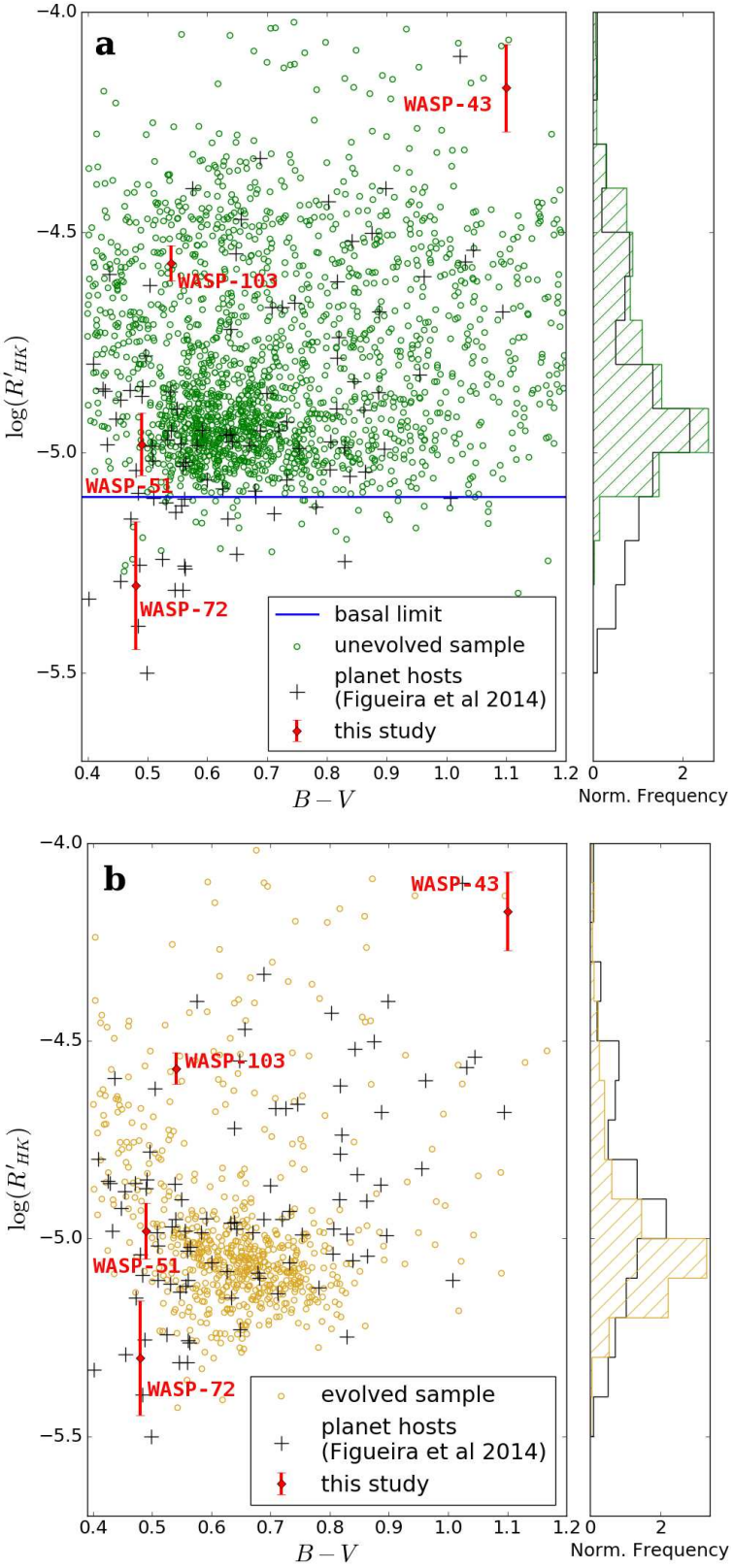

Figure 3.9: $\log(R'_{\mathrm{HKmean}})$ distribution for field stars compared to archival data on planet hosts and our measurements, using **(a)** unevolved, main sequence stars less than 0.45 mag above the average main sequence and **(b)** evolved field stars between 0.45 mag and 2 mag above the average main sequence. Note that bimodality in the distributions is expected due to the Vaughan-Preston gap (see e.g. distributions in Wright 2004, Gray et al. 2006).

The comparisons shown in Figure 3.9 indicate that transiting planet hosts have a greater spread in $\log(R'_{\rm HK})$ than field stars, a statement which remains true whether one considers main sequence or evolved stars. I performed Anderson-Darling tests (Stephens 1974), comparing the $\log(R'_{\rm HK})$ distribution of the planet host sample with our main sequence, evolved, and combined field star samples respectively. The resulting Anderson-Darling statistic values of 27.7, 8.2 and 8.6 correspond to very low probabilities ($3\times10^{-5}, 4\times10^{-4}, 3\times10^{-4}$) that the samples are drawn from the same parent distribution in each case. I find that the Figueira et al. (2014) dataset contains 22 planet hosts below the basal chromospheric emission limit (24 % of the sample). Activity depression is widespread in the known population of transiting planet hosts. Only 2% of our unevolved sample and 9% of our combined field star sample show such an unusual activity level. Fig. 3.10 identifies all planet hosts below the basal limit, more than doubling the number previously identified (Fossati et al. 2013). The horizontal basal envelope which is fairly clear in the main sequence field star sample in Fig. 3.9(a) becomes a diagonal envelope in Fig. 3.9(b), with higher $\log(R'_{\rm HK}) \sim -5.0$ for the bluest evolved stars and lower $\log(R'_{\rm HK}) \sim -5.2$ for the redder evolved stars (c.f. Mittag et al. 2013). WASP-72 and the outliers identified by Fossati et al. (2013) appear even more anomalous in the context of Fig. 3.9(b) than they do in the context of Fig. 3.9(a).

Anomalously low $\log(R'_{\rm HK})$ is seen for Kepler-25 and Kepler-68: multi-planet systems each hosting 2 low-mass (0.02 -0.08 $M_{\rm J}$), short period companions. Tidal influence of these planets on their parent stars is negligible. Therefore, a tidal stellar activity suppression mechanism as proposed for WASP-18 by Miller et al. (2012) can certainly be excluded in these cases. Note that all anomalous planet hosts discussed fall below the chromospheric basal limit, but above $R_{\rm phot}$, the purely photospheric contribution to the Ca II H & K bandpasses. This is shown in Figure 3.11, where I plot $R_{\rm phot}$ as defined by Noyes et al. (1984) and the $R_{\rm HK} = R_{\rm phot} + R'_{\rm HK}$ values for each star (Equations 1.7, 1.8).

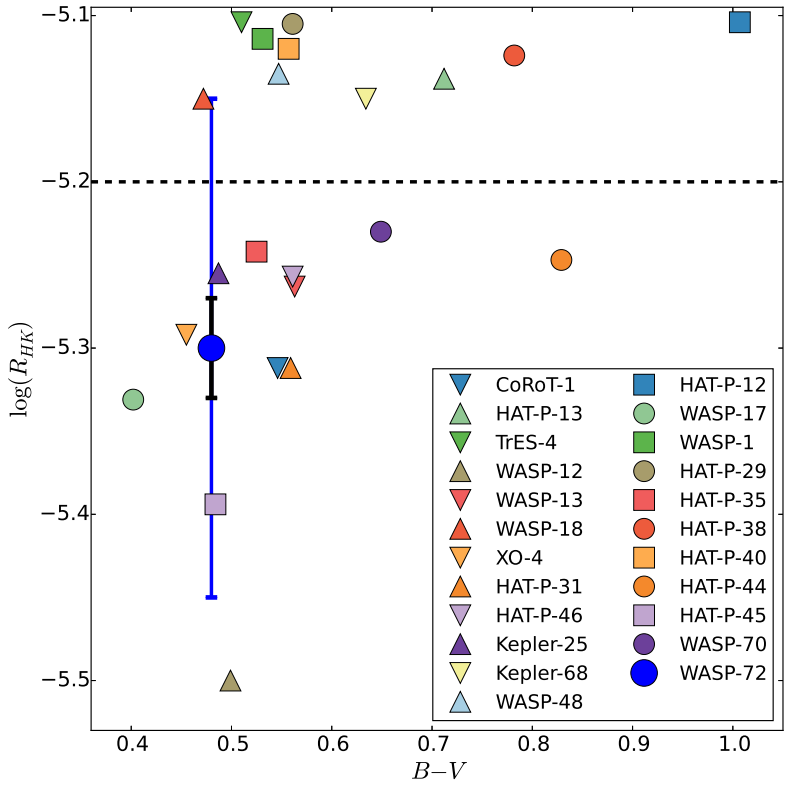

Figure 3.10: Short period planet hosts from Figueira et al. (2014) with $\log(R'_{\rm HK})$ below the main sequence basal limit (-5.1) and our measurement of WASP-72. For reference, I show the uncertainty with (blue) and without (black) the uncertainty in our calibration to the Mount Wilson system (Equation 3.2). Stars with $|\gamma - {\rm V_{ISM}}| < 15$ km s^{-1} (triangles) may be depressed below $\log(R'_{\rm HK}) = -5.2$ by ISM absorption.

Potential ISM contribution to $R_{\rm HK}$ depression

Transiting planet hosts are generally more distant than field stars with $\log({\rm R'_{HK}})$ measurements. Thus an alternative explanation for $\log(R'_{\rm HK})$ depression in planet hosts is absorption in the ISM along the line of sight. In the case of WASP-12, this is unlikely (Fossati et al. 2013), but Fossati et al. (2015c) found that the ISM explanation is viable for WASP-13, while circumstellar absorption could also be present. To assess potential interstellar absorption contributions for the planet hosts of interest, I calculated the difference between the systems' radial velocities, γ, and the velocities of known local ISM clouds, ${\rm V_{ISM}}$ (Table 3.4). I used the Redfield & Linsky (2008) ISM model[10], identifying ISM

[10]http://lism.wesleyan.edu/LISMdynamics.html

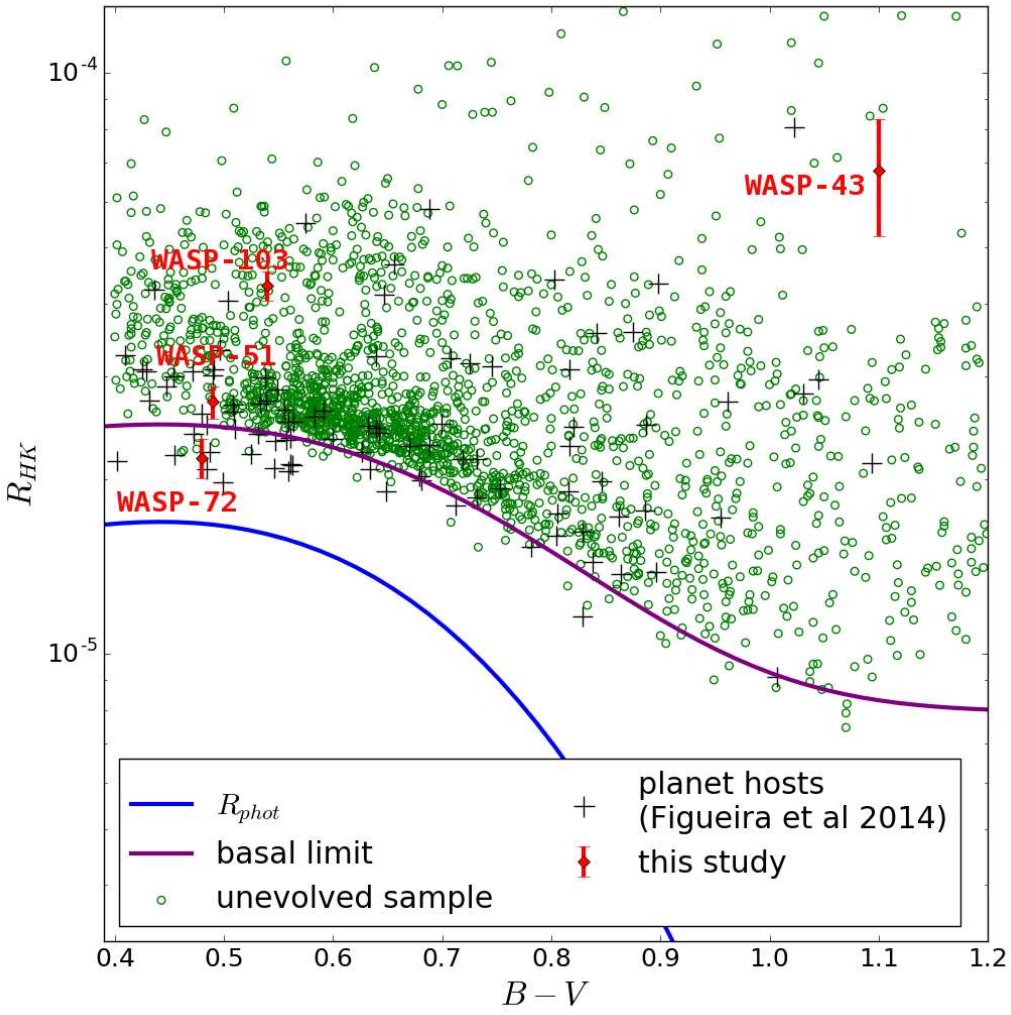

Figure 3.11: $R_{\rm HK}$ distribution for unevolved field stars and planet hosts, compared to the chromospheric basal limit and $R_{\rm phot}$, the photospheric contribution to $R_{\rm HK}$.

clouds traversed along the relevant lines of sight and calculating the clouds' projected radial velocities. Interstellar absorption can depress $\log(R'_{\rm HK})$ values if the velocity differences are within the Ca II H & K core bandpass widths (± 86 km/s). The effect is much more pronounced for smaller $|\gamma - {\rm V_{ISM}}|$, amplified by the triangular H and K bandpass weighting (Fossati et al. 2017). I used the tool[11] developed in Fossati et al. (2017) for the following, approximate estimates of interstellar contributions to $\log(R'_{\rm HK})$ depression in our sample. If we assume typical ISM column densities of $\log(N_{\rm CaII}\ [{\rm cm}^{-2}]) = 12$ (Welsh et al. 2010) for

[11] http://geco.oeaw.ac.at/software.html

the relevant stellar distances (100-500 pc) and conservatively assume a star has intrinsic chromospheric emission exactly at the basal level, our calculations show that $|\gamma - V_{ISM}| \lesssim 15$ km s^{-1} is required to cause $\log(R'_{HK})$ values lower than -5.2. I indicate $\log(R'_{HK}) = -5.2$ in Fig. 3.10, and differentiate objects with $|\gamma - V_{ISM}| \lesssim 15$ km s^{-1} from the remainder. The objects with $\log(R'_{HK}) < -5.2$ and $|\gamma - V_{ISM}| > 15$ km s^{-1} (WASP-17, WASP-70, WASP-72, HAT-P-35, HAT-P-44 & HAT-P-45) are those most likely to be exhibiting significant circumstellar absorption of the Ca II H & K line cores. Using the assumptions outlined above and the $|\gamma - V_{ISM}|$ values in Table 3.4, I calculated values of the interstellar contribution to the $\log(R'_{HK})$ depression for individual stars to be $\sim$ 0.03 - 0.12 dex. Note that we lack data on column densities at different $|\gamma - V_{ISM}|$ values. To be conservative, I always assumed the entire absorption takes place at the smallest $|\gamma - V_{ISM}|$ values in Table 3.4, resulting in upper limit estimates for the effect on $\log(R'_{HK})$.

Table 3.4 and the line in Fig. 3.10 is only indicative of potential ISM contributions from known local clouds, mapped at low spatial resolution. A detailed case by case analysis (see Fossati et al. 2013 and Fossati et al. 2015c) is needed for all anomalous planet hosts to assess the ISM versus circumstellar absorption contributions. This needs to take into account the likely intrinsic activity levels in all cases, and is beyond the scope of the current work.

Table 3.4: Differences between the stellar radial velocities (γ) and projected velocities of known ISM clouds (Redfield & Linsky 2008), traversed along the line-of-sight to each planet host below the chromospheric basal limit, and WASP-51. γ values were taken from the Exoplanet Orbit Database and the individual planet discovery papers.

System	γ [km s^{-1}]	$\lvert\gamma - V_{ISM}\rvert$ [km s^{-1}]	ISM clouds
CoRoT-1	23.8	3.2, 0.6	LIC, Aur
HAT-P-12	-40.6	34.7	NGP
HAT-P-13	14.8	0.3	LIC
HAT-P-29	-21.7	36.8	LIC
HAT-P-31	-2.4	14.7, 21.2, 26.5	LIC, Mic, Oph
HAT-P-35	41.0	24.2, 20.9	LIC, Aur
HAT-P-38	-19.7	37.4, 33.2	LIC, Hya
HAT-P-40	-25.0		none
HAT-P-44	-33.5	27.5	NGP
HAT-P-45	23.9	55.3, 50.6, 63.5	Oph, G, Aql
HAT-P-46	-20.9	10.0, 6.5, 21.5	Oph, G, Aql
Kepler-25	-8.5	2.9, 10.5, 5.0	LIC, Mic, G
Kepler-68	-20.9	13.2, 5.8	LIC, Mic
TrES-4	-15.9	1.8, 4.3	LIC, Mic
WASP-1	-13.4	22.6	LIC
WASP-12	18.9	3.2	LIC
WASP-13	9.9	2.7	LIC
WASP-17	-49.3	21.0	G
WASP-18	2.8	2.0, 27.3, 8.4	LIC, Dor, Cet
WASP-48	-19.7	14.2	LIC
WASP-70	-65.4	43.6	Mic
WASP-72	35.9	23.4, 17.8	LIC, G
XO-4	1.6	14.2	LIC
WASP-51	44.7	28.0, 24.8	LIC, Aur

The Roche lobes and surface gravities of the Hot Jupiter planets hosted by our target stars

It is well-known that the mass loss rate expected from a Hot Jupiter planet will depend on the irradiation from the host star (Lammer et al. 2003), but it will also depend on the depth of the gravitational potential well to be overcome. For close-in giant planets, the Roche lobe can be substantially filled, which can have a significant effect on the structure of the planet (Burton et al. 2014) and must reduce the energy required for escape, compared to that for escape from an isolated planet (Erkaev et al. 2007, Burton et al. 2014). In the extreme case, that of a completely Roche-lobe filling planet, planetary material would flow across the inner Lagrangian point towards the star with no energy barrier to overcome. Note that (i) the Roche geometry assumes both the planet and the star are sufficiently centrally condensed to be treated as point sources of mass, an assumption which is not as well satisfied for giant planets as it is for stars; (ii) only the gravitational and centripetal forces are included. Several other phenomena are neglected: radiation pressure, magnetic forces on charged particles and Coriolis forces on moving particles. Nonetheless, the degree of Roche-lobe filling is a useful parameter to consider for transiting planets, i.e. where there is an empirical measurement of the planetary size. We calculated the Roche geometry for the target systems, using parameters from TEPCAT as adopted throughout this paper. We corrected the empirical radius for the projection effect of a non-spherical planet viewed at orientation $i \neq 90^{\circ}$, to give a self-consistent 3-D Roche geometry. See Busuttil (2017) for details on the Roche calculations, and Table 3.1 for the relevant inclinations. The results of our calculations are shown in Fig 3.1, which gives an immediate feeling for the degree of Roche-lobe filling in each case. We calculated the projection-corrected volume of the planet and that of the corresponding Roche lobe in each case. The percentage filling is 11.82% , 5.83%, 3.92 % and 44.62% for

WASP-43b, WASP-51b, WASP-72b and WASP-103b respectively. For reference, we calculate the filling factors of WASP-12b, HD 189773b and HD 209458 to be 60.54%, 3.72% and 4.02% respectively.

Figure 3.12 shows our results in the context of the observed correlations between stellar activity and planetary surface gravity for a subset of transiting planet-hosts, as described in Fossati et al. (2015b). Note that membership of the active and inactive subsets is assigned using an a posteriori probability criterion, and not with a specific activity cut-off level. Broadly, our values match the archival activity distribution of Fossati et al. (2015b); I will discuss individual targets and implications in Section 3.4.2.

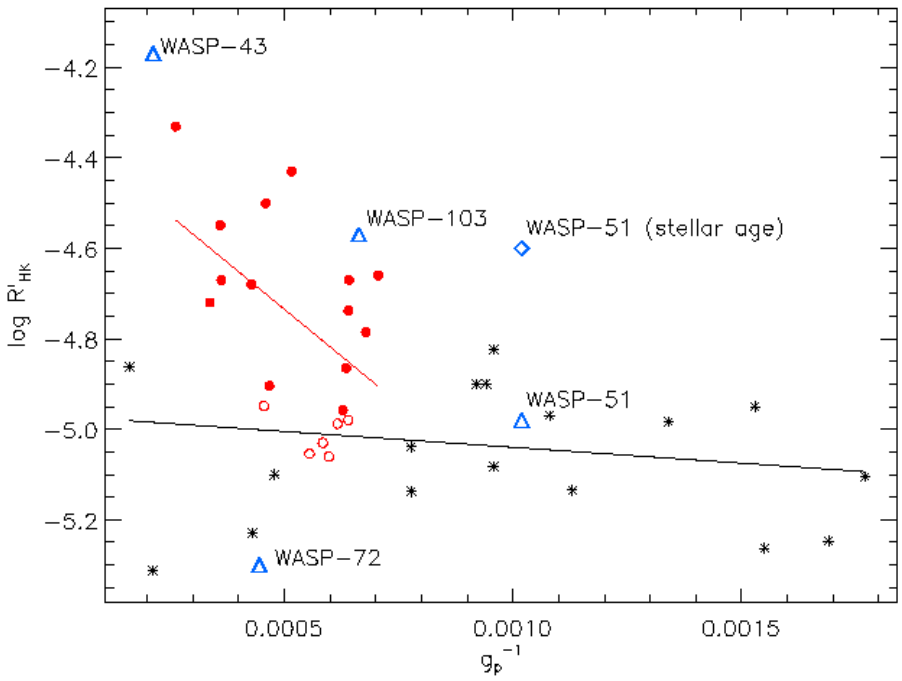

Figure 3.12: Updated plot of $\log(R'_{\rm HK})$ against the inverse of planetary surface gravity ($g_{\rm p}$) from Fossati et al. (2015b), with their mixture model correlation fits for active (red) and inactive stars (black). Stars with ambiguous categorisation are plotted as open circles. The measurements from our study (triangles) are labelled. For reference, I include the approximate activity level for WASP-51 expected from the stellar age (diamond); see section 3.4.2 for details.

3.4.2 Findings for individual systems

WASP-43

WASP-43b is one of the few HJs discovered around a relatively late, K-type star. Due the systems' proximity, at $d \sim 80$ parsec, and favorable star-planet parameters it has been extensively studied. Photometric and spectroscopic investigations both from the ground (e.g. Chen et al. 2014, Murgas et al. 2014) and space (Kreidberg et al. 2014, Stevenson et al. 2014, Wong et al. 2015) have yielded detailed insights into the planet's atmosphere. The very short orbital period, 0.81 d, does not lead to irradiation at the levels experienced by WASP-103b and WASP-72 (c.f. Table 3.1), due to WASP-43b's late type host ($T_{\rm eff}$ = 4520 K). The planet experiences nearly identical irradiation levels as HD 209458b, with a period of 3.5 d (Kataria et al. 2015).

WASP-43 has a rotation period of 15.6 d, i.e. a gyrochronological age of 300-600 Myr (Hellier et al. 2011). This is consistent with its X-ray emission (Czesla et al. 2013), but tidal spin-up could affect both, obscuring the true system age. Our $\log(R'_{\rm HK})$ value places WASP-43 at the activity level seen for very young (< 130 Myr; MH08) members of the Pleiades cluster and Sco-Cen association in Figure 3.9. WASP-43 lies well above the Hyades, $\sim$ 625 Myr, activity level at this spectral type; see Figure 2.11. The $\log(R'_{\rm HK})$-age relation of MH08 gives a very low age (Table 3.3), inconsistent with WASP-43's gyrochronological age. While the scatter around this relation is large for young stars (at the 60 % level in age, MH08), our $\log(R'_{\rm HK})$ derived value still clearly disagrees with the gyrochronological age. The activity value is not significantly affected by erroneous extrapolation of our calibration to large S-values. To verify this, we compared our spectrum of WASP-43 with an archival observation: from the Pace (2013) catalogue, I selected a comparison star with basic stellar parameters close to WASP-43's values. The star with the highest activity value that meets these criteria is HD 86856, with $\log(R'_{\rm HK})$ = -4.37,

measured from a HIRES spectrum. We degraded this archival spectrum to our RSS resolution; the comparison confirmed that WASP-43's core emission is higher by approximately the amount expected from our $\log(R'_{HK})$ value.

MH08 found a tight correlation between stellar X-ray emission, parameterised by $\log(R_X)$, and $\log(R'_{HK})$. Using WASP-43's $\log(R_X) = -4.98 \pm 0.23$ (Czesla et al. 2013), equation A1 in MH08 predicts $\log(R'_{HK}) = -4.56 \pm 0.07$. Our measured value, $\log(R'_{HK}) = -4.17 \pm 0.10$ constitutes a clear outlier with respect to the MH08 correlation of X-ray and Ca II H & K emission in their sample of ~ 200 stars. To highlight this, I replotted Figure 15 of MH08 with our WASP-43 measurement in Fig. 3.13.

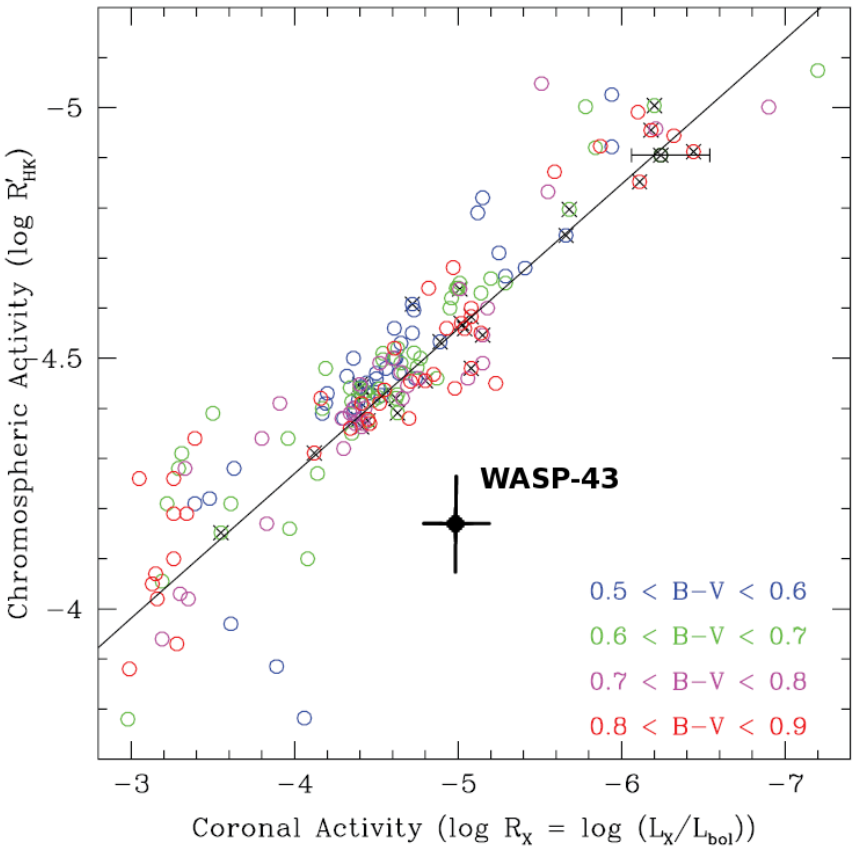

Figure 3.13: Comparison of our WASP-43 $\log(R'_{HK})$ measurement and its $\log(R_X)$ value (Czesla et al. 2013) in the context Figure 15 of MH08.

The X-ray activity level agrees with the stellar rotation period, and our $\log(R'_{HK})$ measurement is anomalously high. It is possible that this is due to a stellar flare during the exposure. If the X-ray derived $\log(R'_{HK})$ level represents

a quiescient baseline, this implies an increase in Ca II H & K emission by a factor of 2 to 3, given our uncertainties in $\log(R'_{\rm HK})$. Such short-term variability has been observed for the young, rapidly rotating K5V star BD+201790, which hosts a disputed HJ (Hernán-Obispo et al. 2010; Hernán-Obispo et al. 2015). However, stars similar to WASP-43 with $4000\,{\rm K} < T_{\rm eff} < 5000\,{\rm K}$; and $10{\rm d} < P_{\rm rot} < 20{\rm d}$ have flare frequencies of 0.04 d^{-1} - 0.21 d^{-1} with durations of 0.1 - 15 hours (Balona 2015). This implies a < 14% chance of seeing a flare during our 0.56 hour exposure even at the most favourable flare frequencies and durations.

Our measurement was taken at WASP-43 b orbital phase 0.66, whereas the X-ray data from Czesla et al. (2013) covers the secondary eclipse of the planet. Shkolnik et al. (2008) observed clear Ca II H & K emission increases for the HJ host HD 179949 at phases 0.6 - 1 and attributed this to SPI. If WASP-43 shows similar behaviour, our data could be affected by SPI, whereas the X-ray data is not. The putative strong flaring event during our exposure could have been induced by SPI increasing the flaring frequency above that of stars without HJ companions. Only mixed evidence exists for Ca II H & K emission increases due to SPI at the few percent level (Miller et al. 2015). The doubling or tripling implied by our measurement is unique, and requires confirmation.

Figure 3.14 shows all planets from the Exoplanet Orbit Database with $0.1\,M_{\rm J} < M_{\rm p} \sin i < 13\,M_{\rm J}$ and $a < 0.1$ AU, the Hot Jupiters observed in this paper and several systems where evidence for SPI has been published. I plotted the tidal SPI proxy $h_{\rm tide}/h_{\rm scale}$, following Cuntz et al. (2000), where $h_{\rm tide}$ is the height of the tidal bulge raised on the star, and $h_{\rm scale}$ is photospheric scale height (see also e.g. Poppenhaeger & Wolk 2014; Pillitteri et al. 2014). In the absence of detailed theoretical modelling and magnetic field maps, simple scaling proxies based on known system properties are the only available guide to assess putative magnetic SPI strength. Ideally, the maximum power released by magnetic reconnection would be calculated analytically following Lanza (2013). A meaningful estimate requires measurements of the stellar field strength and

configuration, which are not available for our targets. It is worth noting that Miller et al. (2015) use $M_{\rm p}{\rm sin}i/a^2$ as a magnetic SPI strength proxy, highlighting potential activity enhancement in "extreme" systems with $M_{\rm p}$ sin$i/a^2 > 450$ (see their Fig. 7.) WASP-103 and WASP-43 have $M_{\rm p}$ sini/a^2 values 8 and 20 times larger than this threshold. However, this is not an ideal proxy to differentiate magnetic from tidal SPI effects: $M_{\rm p}{\rm sin}i$ strongly correlates with $h_{\rm tide}/h_{\rm scale}$ due to the $M_{\rm p}{\rm sin}i$ dependence. It is also questionable to assume all close-in planets of interest orbit in the regime where the stellar field strength falls of as a^{-2}, since this will change with the geometry of the fields (see e.g. Lanza 2013). I chose to simply plot the semi-major axis in Figure 3.14 as an indicator of magnetic interaction strength. There is insufficient information about the stellar field environments for the vast majority of exoplanets to differentiate them into subsets of other parameterisations. Both proxies plotted are exceptionally favorable for the WASP-43 system. The alternative simple proxies used in the literature, such as $M_{\rm p}{\rm sin}i/a^2$, $M_{\rm p}{\rm sin}i/a$ and $M_{\rm p}{\rm sin}i/P$ (Miller et al. 2015; Poppenhaeger et al. 2010; Shkolnik et al. 2008) also all indicate SPI is particularly likely for WASP-43. Note that WASP-43b and WASP-103 are the Hot Jupiters with the shortest and 4th shortest semi-major axis discovered to date. Since our activity value is anomalously high with respect to the stellar rotation period, a tidal explanation (invoking enhanced rotationally-driven emission) is disfavored for WASP-43.

There is no evidence for circumstellar absorption in accordance with the Salz et al. (2016) finding that WASP-43 b should be stable against mass loss. Time-variable magnetic SPI at an unprecedented level could be an explanation for our high value of $\log(R'_{\rm HK})$. To distinguish this explanation from "normal", flare activity, phase-resolved monitoring of X-ray and/or optical activity indicators of WASP-43 is needed.

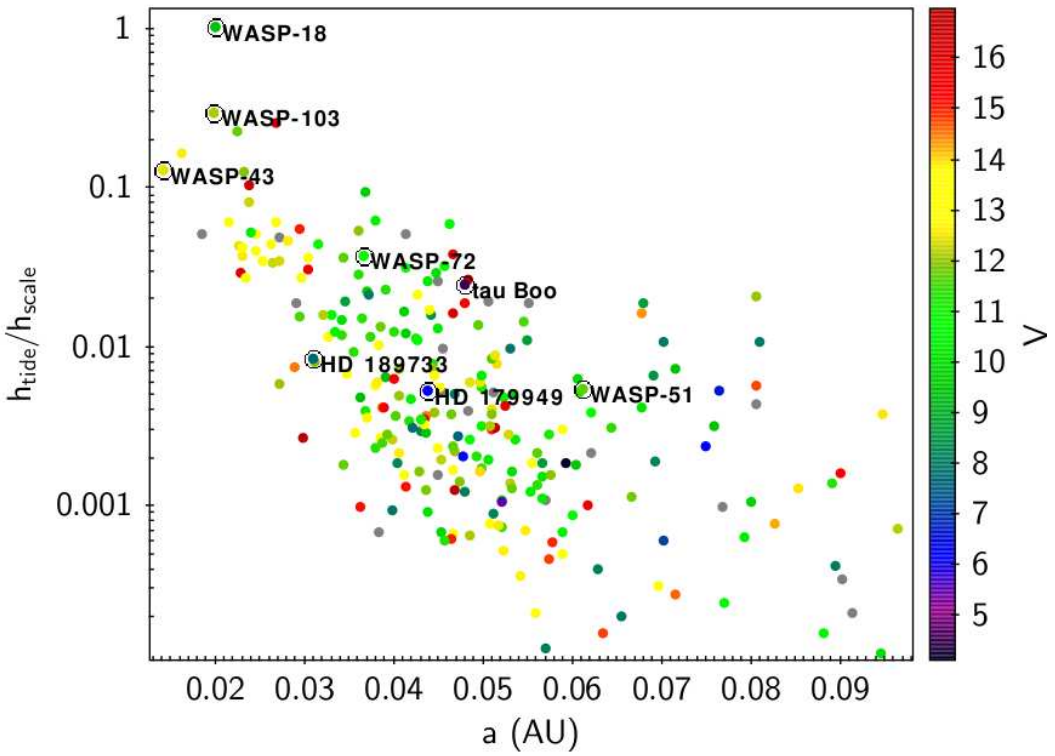

Figure 3.14: SPI interaction strength proxies for planets with $M_p \sin i$ = 0.1-13 M_J and a < 0.1 AU. The visual magnitude is shown as a guide to the observability of the systems. Planets observed in this study and key systems from the literature are labelled. A null-result for SPI in the particularly extreme WASP-18 system was found by Miller et al. (2012), and attributed to the very low intrinsic activity level of this F6V star; see X-ray nondetection in Pillitteri et al. (2014). Evidence for weak SPI effects have been reported for HD 179949 (Shkolnik et al. 2003), HD 189733 (Poppenhaeger & Wolk 2014) and tau Boo (Walker et al. 2008).

WASP-51/HAT-P-30

This planet was discovered independently in 2011 by the HATNET and SWASP surveys and is therefore referred to as either HAT-P-30 or WASP-51. It is a typical spin-orbit misaligned HJ (Johnson et al. 2011; Enoch et al. 2011) and has received limited dedicated follow-up to date. Our activity value is above the basal level, and corresponds to a MH08 age of 6.2 ± 1.3 Gyr. Note that MH08 reported a scatter around their activity-age relation at the 30% level for ages older than Pleiades (130 Myr). The most precise literature ages (0.6-1.8 Gyr; Table 3.3) are based on isochrone dating, using the stellar parameters of Johnson et al. (2011). The high stellar Li abundance is consistent with a ≤ 1 Gyr age

(Enoch et al. 2011). Our $\log(R'_{HK})$ derived age (Table 3.3) is clearly anomalously high. In summary, it seems the true age is ~ 1 Gyr and the normal stellar activity has been depressed, either through the planet tidally suppressing the stellar magnetic dynamo or through absorption of the Ca II H & K emission by circumstellar gas from the planet or by the ISM. The MH08 relation predicts WASP-51 should exhibit $\log(R'_{HK})$ ~ -4.6 at a 1 Gyr age. I plotted this value in Fig. 3.12 to highlight the discrepancy. Such high intrinsic activity could retain the observed, depressed value above the basal limit.

If the intrinsic chromospheric emission is indeed at the $\log(R'_{HK})$ ~ -4.6 level, our measurement implies depression of $\log(R'_{HK})$ by ~ 0.4 dex. This is unlikely to arise from absorption in the ISM: our calculations show that clouds with $|\gamma - V_{ISM}| \lesssim 30$ km s^{-1} and a column density $\log(N_{CaII}\ [\mathrm{cm}^{-2}]) > 16$ would be required. For stars out to 800 pc Welsh et al. (2010) consistently observe $\log(N_{CaII}\ [\mathrm{cm}^{-2}]) < 13$ (typically ~ 12), and WASP-51 lies at ~190 pc (Johnson et al. 2011).

Note that the mean activity level reported in Johnson et al. (2011) using the HIRES spectrograph is $S_{MW} = 0.128 \pm 0.014$. The uncertainty is taken to be 11%, as reported for the HIRES calibration to the Mount Wilson system (Isaacson & Fischer 2010). The difference between this S_{MW} value and our measurement (Table 3.3) is at the 1.5 σ level, i.e. not significant.

The surface gravity of WASP-51b is relatively low (Fig. 3.12). There is a conspicous lack of low surface gravity HJs with high $\log(R'_{HK})$ values (Fig. 3.12). Low surface gravity planets are most prone to prodigous mass loss, so their host stars are the most likely to be surrounded by circumstellar gas shrouds fed by mass loss from the planet. Absorption by these shrouds depresses the stellar Ca II H & K emission, moving points which would otherwise appear in the top right of Fig. 3.12 down towards the bottom right, as indicated for WASP-51. The complete lack of systems in the top right of Fig. 3.12 suggests that such circumstellar gas shrouds may be a common feature of the known low surface

gravity transiting exoplanets. This provides circumstantial evidence in favour of the gas shroud explanation for *apparent* activity suppression being more widespread than ISM absorption or the tidal suppression of the actual stellar activity level. Of course the latter processes may still operate in some range(s) of parameter space.

WASP-103

WASP-103b is an extreme HJ in terms of irradiation level and expected tidal effects (Gillon et al. 2014). Orbital decay is predicted to be measurable over the next decade (Birkby et al. 2014), and it is one of the most tidally distorted planets known, close to tidal disruption (Southworth et al. 2014). Among our 4 targets, it is the most similar system to the extreme case of WASP-12b.

The $S_{\rm RSS}$ RMS variability over our 16 observations of WASP-103 is at the 2 % level, comparable to the photon noise uncertainties. The dispersion in $\log(R'_{\rm HK})$ is 0.02. This is lower than the uncertainties quoted for individual values, since those include additional systematic calibration and $B-V$ uncertainties. I divided all RSS spectra by the spectrum with highest SNR, and confirmed there are no significant changes in the Ca II H & K line core profiles. Unfortunately the data sample a narrow phase range away from transit, so the issue of phase-dependent absorption cannot be addressed. To achieve this goal, I recommend high cadence observations that sample the full phase range, taken over a short baseline to minimise intrinsic stellar activity variation from stellar rotation and spot evolution. SALT is not an ideal telescope for such a program, due to its queue-schedule and visibility constraints.

There are two age estimates for WASP-103 in the literature: fitting to stellar evolution models yielded 1.8 - 6.2 Gyr (Southworth et al. 2014), while taking into account the stellar Li abundance of $\log A$(Li) 2.23 ± 0.13 as an additional constraint gave an estimate of 3 - 5 Gyr (Gillon et al. 2014). Our mean $\log(R'_{\rm HK})$ value of -4.57 ± 0.04 corresponds to an activity-age of 950 ± 230 Myr using the

relation of MH08. The significance of the age difference is only high ($\geq 3\sigma$) with respect to the Gillon et al. (2014) age estimate. Improved stellar parameters are needed for a better constrained independent age assessment (c.f. Table 6 in Southworth et al. 2014). By analogy with WASP-12 (Fossati et al. 2010a; Haswell et al. 2012) we expected depressed activity due to absorption by circumstellar gas from the extreme HJ planet. Either this is masked by SPI or the planetary mass loss rate may be unexpectedly low. Modelling of atmospheric escape, as for WASP-43, would be useful to investigate this possibility. Figure 3.14 suggests SPI enhancements are plausible. WASP-103 has the second-highest tidal interaction proxy value $h_{\rm tide}/h_{\rm scale}$ of all known exoplanets. The value is a factor of three lower than that of the extreme WASP-18 system, for which the possibility of tidal activity suppression rather than enhancement has been invoked (Pillitteri et al. 2014).

WASP-72

WASP-72 is among the most highly irradiated Hot Jupiters known (Gillon et al. 2013), experiencing approximately 60% of the WASP-103b irradiation level. There have been no follow-up studies of this planet to date. Our measured activity value falls below the basal level, see Figure 3.9 and 3.10. While the uncertainty on $\log(R'_{\rm HK})$ is relatively large, WASP-72 certainly falls amongst the lowest activity outliers for both main sequence and evolved stars of its spectral type. Figure 3.10 shows the substantial uncertainty contribution from the calibration to the Mount Wilson system. Note that intercomparing activity values measured with the same instrument setup does not suffer this systematic uncertainty.

The MH08 activity age implied by our measurement is 11 ± 0.8 Gyr, while stellar evolution codes give 2.6 - 3.8 Gyr (Gillon et al. 2013). This is the only literature age estimate; it strengthens the case for a depressed activity value. From MH08, the expected $\log(R'_{\rm HK})$ over this age range is approximately -4.8 to

-4.9. In the context of Figures 3.12 and 3.9, WASP-72 may therefore be a system where relatively weak emission from an inactive, ~3 Gyr old star is absorbed sufficiently to appear below the basal limit, analogously to other close-in planet systems (Lanza 2014; Fossati et al. 2015b).

With currently available data for the WASP-72 system, circumstellar absorption from planetary mass loss cannot be disentangled from absorption in the ISM. However, my estimates in Section 3.4.1 indicate the known ISM clouds along the line of sight (Table 3.4) cannot entirely explain the anomalously low value even if the intrinsic stellar emission is exactly at the basal level.

3.5 Summary

We calibrated the RSS at SALT to measure chromospheric activity on the Mount Wilson system. We used this to measure the activity of four HJ host stars. Fossati et al. (2013) highlighted the anomalously low $\log(R'_{\rm HK})$ values for WASP-12 and five other planet hosts included in Knutson et al. (2010). I revisited this with the significantly extended dataset of Figueira et al. (2014), finding that 24 % of the sample (22 hosts) show these anomalies including two low mass, multi-planet systems and WASP-72 (Figs. 3.9 and 3.10).

WASP-43 shows anomalously high activity, while WASP-51/HAT-P-30 has an anomalously low $\log(R'_{\rm HK})$ for its age. There may be (at least) two processes operating in close-in HJ systems which affect the observed $\log(R'_{\rm HK})$ values. Anomalously high $\log(R'_{\rm HK})$ values may be attributed to SPI (Cuntz et al. 2000). The outliers in the bottom left of Fig. 3.9 can be attributed to diffuse circumstellar gas lost from the planets which absorbs the stellar chromospheric emission in strong resonance lines (Haswell et al. 2012; Fossati et al. 2013). Absorption by the ISM may also play a role for some of these systems. My estimates suggest the latter cannot explain the anomalous values of WASP-51 and WASP-72. Further investigations should examine any interstellar and circumstellar

absorption contributions, as was done for WASP-12 (Fossati et al. 2013), WASP-18 (Fossati et al. 2013) and WASP-13 (Fossati et al. 2015c), and test whether the intrinsic activity levels of WASP-72 and WASP-51 are indeed normal.

Further measurements are needed to investigate whether WASP-43's anomalously high $\log(R'_{\rm HK})$ is time-dependent. The extreme HJ WASP-103 b also displays higher activity than expected from the system age; improved stellar parameters are needed to test this conclusion. WASP-43 and WASP-103 have exceptionally high predicted SPI levels from the simple scaling laws illustrated in Fig. 3.14. No evidence for stellar activity supression in WASP-103 was found. By analogy with WASP-12b one might have expected to find significant absorption due to planetary mass loss.

Activity can appear depressed by circumstellar absorption of the stellar Ca II H & K flux. SPI can potentially enhance the normal Ca II H & K emission. In some systems, including WASP-103, both mechanisms may operate simultaneously. Such systems would not necessarily have obviously anomalous $\log(R'_{\rm HK})$ activity ages. We recommend that the $\log(R'_{\rm HK})$ index is never used as an age indicator for stars which host close-in planets: it is unreliable for this purpose. Disentangling the mechanisms to the extent of being able to predict which one dominates will require extensive uniform observational work, covering a range of stellar spectral types, ages, planet masses, planet radii, Roche lobe radii, orbital separations, and measurements of the stellar magnetic field configurations.

Details on collaborative contributions

My supervisor C.A. Haswell initiated and designed the RSS project and led the observing proposals for SALT telescope time (including selection of a subset of targets and the instrument mode used), carried out significant edits of the paper manuscript that this chapter builds on, and originated some of the points in the conclusions. Observations were carried out by the local SALT astronomer

team. I implemented and tested all data reduction steps on two initial sets of observations. All subsequent reductions for the remaining targets were carried out by G.D. Smith, under my supervision. L. Fossati provided valuable input for the interstellar absorption analysis and discussion, including Figure 3.12. As referenced, I made use of the analytical tool he developed in this context, for which I provided significant input. R. Busuttil performed the Roche Lobe calculations and created Figure 3.1. J.R. Barnes reduced and analysed the archival HIRES spectrum I used as a comparison for WASP-43 in Section 3.4.2). All co-authors of the paper that forms the basis of this thesis chapter (C.A. Haswell, G.D. Smith, L. Fossati, J.R. Barnes, R. Busuttil, J.S. Jenkins) provided feedback on the manuscript.

Chapter 4

Radial velocity planet search around targets of interest: Methods

In this chapter, Section 4.1 discusses the instrumentation, data reduction and analysis required for precision RV measurements. In Section 4.2, I summarise astrophysical RV signals with non-planetary origins. Section 4.3 gives an overview of the spectrographs used for our RV planet search, and describes important RV corrections I implemented and systematic effects I analysed. The feasibility of our RV programme is presented in Section 4.4 in the form of simulated detection limits. Finally, Section 4.5 describes the methods I used to detect planetary signals in the RV timeseries.

4.1 Instrumentation and data reduction for precision radial velocities

Measurements of precision radial velocities via the Doppler effect are obtained from two fundamentally different methods, using either an iodine gas cell or a stabilised spectrograph with simultaneous wavelength reference. The iodine method can be employed on general-purpose spectrographs without the requirement of a very stable instrumental profile. Starlight passes through the gas cell at the instrument entrance, which superimposes a dense forest of iodine lines on the stellar spectrum. Crucially, all iodine and stellar absorption lines suffer the same instrumental distortions. Butler et al. (1996) pioneered this method, enabling many early RV planet discoveries.

All RV data acquired for this thesis are derived from the simultaneous reference approach, which I will discuss in detail below. With both methods, precision of order 1 $\mathrm{m\,s^{-1}}$ has been demonstrated (Fischer et al. 2016). This is a spectacular achievement, corresponding to Doppler shifts on the order of $1{\times}10^{-5}$ Å, or 0.001 pixels on the detector of a high resolution ($R \sim 100{,}000$) instrument (Bouchy et al. 2009a). Such a feat is only possible by exploiting high resolution ($R \sim 100{,}000$) spectra covering thousands of stellar absorption lines over a wide wavelength range. Echelle spectrographs are the natural type of instruments enabling such measurements. Due to their cross-dispersed optical path layout, spectra can be imaged onto detectors with sensible dimensions: 1-3 standard CCDs are typically needed to cover the full wavelength range of typically ~3800 to 7000 Å (see below).

The Doppler information content of spectra is constrained by stellar parameters: a large number of sharp, deep lines provide the most accurate RVs. The spectra of early type ($T_{\mathrm{eff}} \gtrsim 10\ 000$ K) and rapidly rotating stars show few, broadened lines providing very limited Doppler information. RV planet detection can become impossible for these stars. For spectral types earlier than

B8, the achievable detection limits do not reach the planetary mass regime even at the shortest orbital periods (Lagrange et al. 2009). In our program, several of our targets (mostly F-type stars) are moderately fast rotators with $v \sin i$ up to ~ 15 km s^{-1}. The effect of rotational line broadening is significant: photon-limited RV uncertainties are proportional to $(v \sin i)^{1.5}$ for early (F-type) stars (Galland et al. 2005). At SNR = 100, photon-limited uncertainties below 1 m s^{-1} are achievable with a state-of-the-art instrument for slowly rotating stars ($v \sin i \lesssim 2$ km s^{-1}). This single-measurement precision degrades to 7 m s^{-1} at $v \sin i = 15$ km s^{-1} (Galland et al. 2005). Increased stellar activity levels for rapidly rotating, active stars (c.f. Section 4.2) can cause RV variability well above these measurement uncertainties.

4.1.1 Simultaneous reference method

This approach relies on an extremely stable instrumental profile on all timescales throughout many years of operations. The spectrographs used are a) fibre-fed, with efficient scrambling of the starlight in the fibre link and b) environmentally stabilised, minimising pressure and temperature variation. Scrambling of the stellar light mitigates illumination variations at the fibre entrance and achieves a homogeneous, time invariant light distribution at the fibre exit within the instrument. Therefore the instrumental profile does not vary over time. Scrambling is partly an intrinsic property of optical fibres and is augmented via a so-called "double-scrambler", using two separate fibres linked via lenses (Brown et al. 1991).

Even for the best instruments, nanometer-scale flexures in the optical system cannot be avoided, causing instrumental RV drift over time (Pepe et al. 2002b). This is tracked with a (simultaneous) wavelength reference source.

Thorium-Argon (ThAr) lamps are typically used due to their large number of emission lines across the entire wavelength range of interest. They enable drift measurements with photon-noise precision of a few cm s^{-1} (e.g. Bouchy et al. 2012). A second optical fibre is used to record the ThAr spectrum - on the detector the stellar and ThAr spectra are parallel to each other in each echelle order (Fig. 4.1). Obviously, it is crucial that there is negligible differential RV drift between the 2 fibres (hereafter Fibre A and B), which is achieved via near-identical optical components and light paths. A problem of ThAr lamps is their large dynamic range between emission lines. To achieve sufficient SNR overall, the brightest lines invariably saturate and cause "bleeding" along the CCD columns, contaminating the stellar spectrum somewhat (Fig. 4.1). For

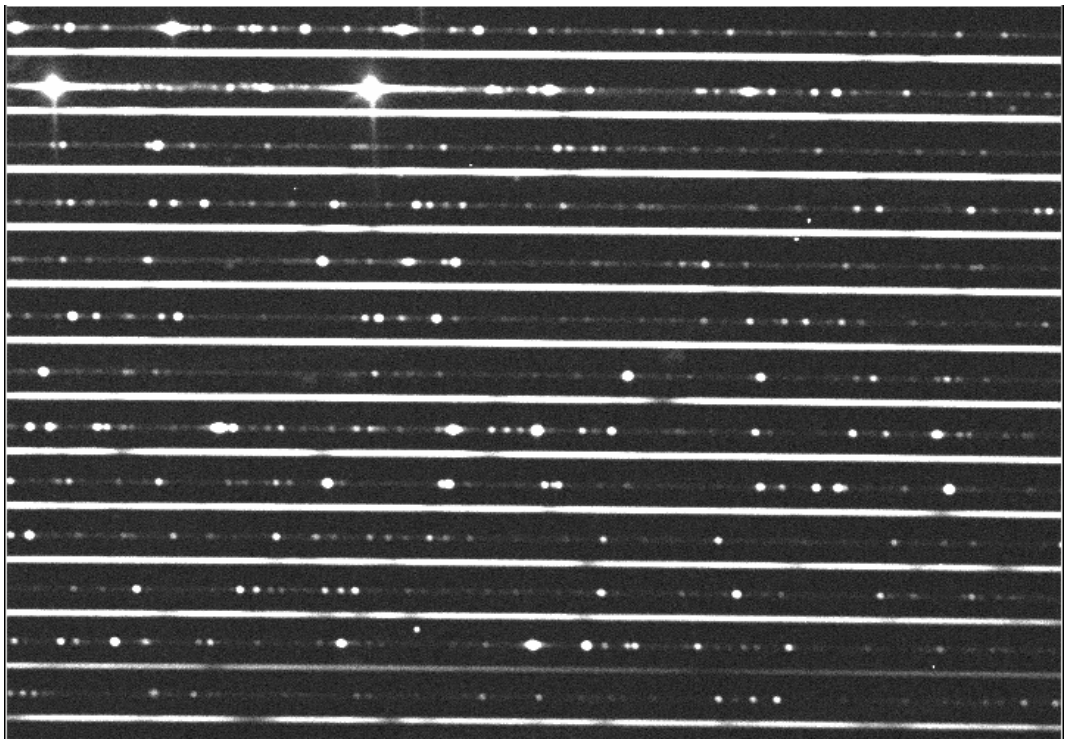

Figure 4.1: A small section of a HARPS CCD frame from our observing programme, illustrating the layout of echelle orders alternating between the stellar spectrum and ThAr emission lines. ThAr line "bleeding" of the brightest, saturated lines is visible in the top left corner.

precision RVs of faint stars, contamination of the stellar spectrum is a serious problem, and ThAr is not used simultaneously. In practise, Fibre B is also used

to record the sky spectrum instead of ThAr emission, which is useful if the sky background is affecting RV measurements (c.f. Section 4.1.2). Since instrumental drifts are usually small, near-simultaneous ThAr exposures before and after stellar observations can often adequately track the drift. I will follow abbreviations used in the literature for the instruments relevant to this thesis (e.g. Díaz et al. 2012): *ThoSimult* mode records ThAr emission with Fibre B, and *objAB* mode records the background sky spectrum.

RV measurements from cross-correlation

Cross-correlation is the original and most widespread approach for measuring RVs for instruments with simultaneous reference. Each extracted, wavelength calibrated echelle order of the stellar spectrum is cross-correlated with a numerical "binary mask". Pioneering work on and a detailed discription of the binary mask approach to cross-correlation can be found in Baranne et al. (1996) and Pepe et al. (2002b). The name is due to the construction of these masks as zero and nonzero values as a function of wavelength, matching a carefully optimised list of theoretical spectral line positions. Weighting of the non-zero values is based on the line depth. A set of masks matching different spectral types can be used. Spectral regions dominated by absorption from the Earth's atmosphere ("telluric lines") are avoided.

The CCF obtained as a result of this process approximates the mean stellar line profile, averaged over the thousands of lines specified by the binary mask. It is approximately Gaussian in shape, so the final RV measurement is the centroid of a Gaussian fitted to the CCF. This measurement is then corrected for instrumental drifts via the RV measured from ThAr on Fibre B.

An important problem for RV measurements are stellar line profile distortions which can arise from several astrophysical phenomena (Section 4.2). To

monitor these, the shape of the CCF can be quantified with several parameters. Historically, the FHWM of the Gaussian fit and the CCF bisector became by far the most widely used metrics (e.g. Figueira et al. 2013). Bisector velocities are simply the midpoints between the blue and red wings of the CCF, evaluated at a set of flux-levels (Fig. 4.2). When using a normalised CCF, these flux levels are expressed in % of the mean line depth. From these bisectors, a "bisector inverse slope" (BIS) is obtained by subtracting the mean bisector velocities between the 90% and 60% flux levels from the mean between the 40% and 10% flux levels (Fig. 4.2). This particular definition is arbitrary, but has been widely adopted for RV exoplanet work following Queloz et al. (2001).

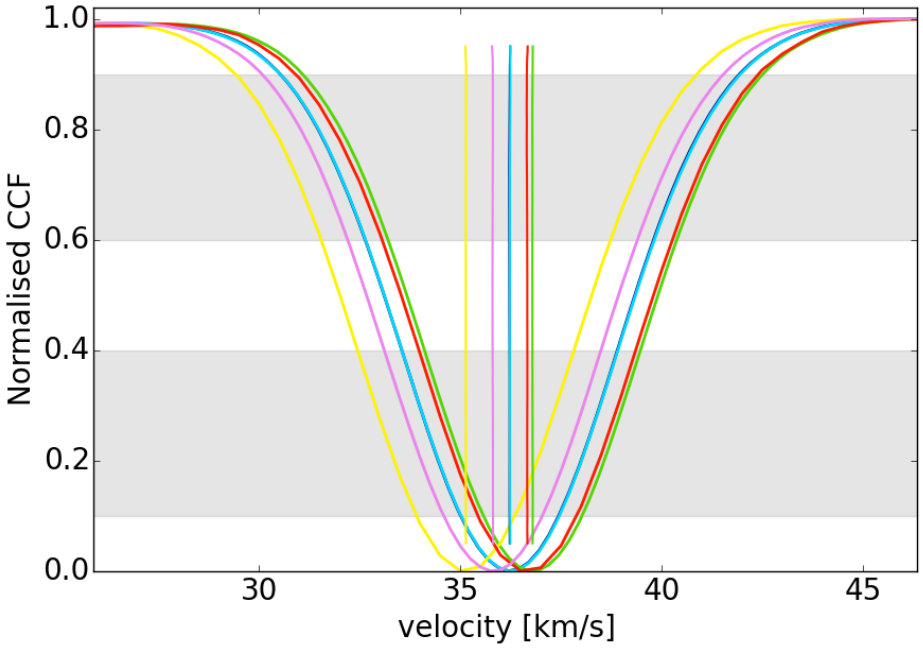

Figure 4.2: CCFs and bisectors for a target in our programme with a Brown Dwarf (BD) companion, observed at 6 epochs. The BD induces a very large reflex RV amplitude of ~ 1.6 km s^{-1}. The corresponding Keplerian parallel shifts of CCF and bisector are clearly visible. Regions used to calculate BIS values are shaded.

4.1.2 Data reduction

Data reduction for echelle spectrographs is a relatively complex task compared to the longslit reduction discussed in Section 3.2. The interested reader is referred to Hall et al. (1994) for a detailed description. To exploit 1 m s^{-1} instrumental precision a very robust, high quality reduction is needed, particularly in the wavelength calibration processes.

For the instruments relevant to this thesis (SOPHIE, HARPS, HARPS-N), sophisticated automated reduction has been implemented and refined over more than a decade[1]. This software, the Data Reduction Suite (DRS), provides near-realtime reduction at the telescopes down to final RV measurements via cross-correlation. Note that the DRS is effectively proprietary software, available only on dedicated workstations at the telescopes. Any (re-)reduction of data can only be performed on these. The main steps performed by the DRS are

- Bias subtraction and flatfielding.
- Tracing of echelle orders using the flatfield frames.
- Optimal extraction of all echelle orders.
- Daily wavelength calibration, using ThAr exposures on both fibers.
- Measurement of the instrumental drift, using the simultaneous ThAr exposures.
- Cross-correlation of each echelle order with binary masks.
- Creation of the weighted average CCF; Gaussian fit and measurements of RV, FHWM & BIS.

[1] www.eso.org/sci/facilities/lasilla/instruments/harps/doc/DRS.pdf

The DRS provides RV uncertainties from three contributions, summed in quadrature: (1) photon and readout noise, (2) wavelength calibration uncertainty, and (3) instrumental drift uncertainty. Photon and readout noise per pixel is propagated to the extracted spectrum, and then to the final weighted CCF. The RV centroid uncertainty on the CCF Gaussian fit is then calculated. The dispersion of ThAr line positions around the wavelength solutions is used to estimate the global uncertainty on the wavelength calibration.

These 3 error sources provide a lower bound on the RV error bars, and are often referred to as internal errors (σ_{RV}). The actual precision achieved even for extremely "well-behaved" RV standard stars is worse than this, due to instrumental systematic effects not captured by the 3 factors and astrophysical variability (Section 4.2). In the literature, this is sometimes dealt with by simply adding uncertainties, so-called RV "jitter", in quadrature to the internal errors before searching for planetary signals (e.g. Brown et al. 2012; Zechmeister et al. 2013; Santos et al. 2014). Such *a priori* accounting of jitter can lead to biased, less reliable recovery of Keplerian signals (Baluev 2009). We use the internal uncertainties and fit a parameter capturing excess RV jitter within our signal search (see Section 4.5 for details). It is important to emphasise that the jitter includes both an instrumental and an astrophysical component.

Uncertainties on BIS and FWHM values are not provided by the DRS. They can be approximated by direct scaling from the internal RV uncertainties: BIS errors are taken to be $2\times\sigma_{RV}$, and FWHM errors are $2.35\times\sigma_{RV}$. See e.g. Zechmeister et al. (2013) for an explanation of the scaling factors.

BERV, the Barycentric Earth Radial Velocity

When measuring the velocity of stars from an observatory on Earth, their variability is dominated by the orbital and rotational motion of Earth at a level of a few $\mathrm{km\,s^{-1}}$. Measured RVs are corrected to the Solar System barycentre frame by evaluating the combined velocity of these 2 effects, the so-called Barycentric Earth Radial Velocity (BERV). Without this correction useful RV observations would be impossible, and fortunately BERV can be calculated to very high precision (c.f. Wright & Eastman 2014). It is established convention to list the Barycentric Julian Date (BJD) as the timestamp of each RV observation, i.e. the signal time of arrival is also referenced to the Solar System barycentre.

Most precision RV spectrographs are equipped with an exposure meter that measures the count-rate throughout an exposure. This information is used to calculate the flux-weighted "effective" midpoint times of the observations, which are governed by time-varying seeing and atmospheric extinction. BERV and BJD can then be accurately evaluated at the effective midpoint, instead of the half-way point. The magnitude of the corresponding change in BERV is completely dominated by Earth rotation and depends on the stellar and observatory coordinates and the observing epoch. The maximum rate of change in BERV, from the equations of simple harmonic motion combined with Earth's radius and rotation period, is $0.034\ \mathrm{m\,s^{-2}}$. A shift of the exposure mid-point by e.g. 200 seconds, which can occur in variable conditions during typical 15-30 min exposure lengths, then translates to a "worst case" change of the BERV by $\sim 7\ \mathrm{m\,s^{-1}}$.

Moonlight contamination

Another important issue for precision RV measurements are high sky background levels. The sky contribution of concern is scattered moonlight, i.e. essentially the solar spectrum. This can be problematic if the radial velocity of the sky spectrum is close to that of the target star. The contaminating spectrum is then quite similar to that of the target, albeit much fainter. The net effect is a subtle asymmetric distortion of the target's line profiles, similar to the spectrum of a spectroscopic binary with very high contrast. This is not detectable in individual lines, but during CCF computation it is sufficient to modify the measured target RV by up to 100 m s^{-1} in extreme cases (Tripathi et al. 2010).

The lunar RV projected along any line of sight can easily be predicted at a given epoch. Its variability is completely dominated by the BERV rotation and orbital periods (Díaz et al. 2012). "Close approaches" between lunar RV and γ can be avoided if the γ values of target stars are known (either from the literature, or after reconnaissance observations). In RV programmes, contamination is either addressed by recording the sky spectrum on Fibre B (Section 4.3.2), or avoided altogether by careful target selection during bright time, preventing RV "close approaches". Tripathi et al. (2010) provide a detailed analysis of contamination effects, including moonlight. The magnitude of the effect depends on the relative RV and SNR of sky and target spectra and on the target's spectral type and FWHM. In practice, a simple "rule of thumb" is recommended by the observing scheduler software in use for HARPS and SOPHIE: when the lunar RV is more than 15 km s^{-1} away from γ, contamination is deemed negligible. In this case, a star can be observed even in bright time. This velocity threshold corresponds to 2 FWHM for solar type stars with low rotation rates ($v \sin i \sim$2 km s^{-1}). For faster rotators, a criterion of 2 FWHM is probably more appropriate than a fixed velocity criterion, based on the work in Cunha et al. (2013).

4.1.3 RVs from the Template Enhanced Radial velocity Re-analysis Application (TERRA)

The CCF method does not exploit the full Doppler information within stellar spectra (Pepe et al. 2002b). TERRA was developed to improve this, and to provide an independent technique to extract RVs from the same datasets analysed via cross-correlation (Anglada-Escudé & Butler 2012). This approach relies on least-squares minimisation of the difference between a high SNR template and individual stellar spectra of a timeseries. Crucially, this template is constructed by co-adding all spectra of the relevant dataset - each star has a unique template. TERRA has been adapted to data from several spectrographs, including HARPS, SOPHIE and HARPS-N. Extracted, wavelength calibrated spectra from the DRS output are the starting point for the template matching approach. TERRA procedures and algorithms are fully described in Anglada-Escudé & Butler (2012). In brief, important steps and features are:

- The weighted sum (over all wavelengths) of the squared difference between template flux and observed flux at each wavelength is minimised.
- This difference is parameterised using terms for the RV shift and the continuum flux normalisation.
- The weighting at each wavelength takes into account poisson noise and a quality factor capturing additional noise.
- The minimisation provides RV shifts and corresponding uncertainties for each echelle order.
- These are then combined into the final RV measurement by a 2-step weighting process.
- Masking of telluric lines and of spectral regions at the edge of echelle orders is used throughout, along with 4σ outlier clipping.

- The simultaneous drift corrections from *objAB* observations and the BERV correction provided by the DRS are applied.

- An initial pass through all the steps described uses the highest SNR spectrum as the template to obtain approximate RV values at all epochs. These RVs are then used to co-add all spectra in the same (barycentric) frame; creating the final high SNR template. The entire template matching procedure as outlined in the previous bullet points is then repeated with this final template.

- Note the perspective (secular) acceleration effect, relevant for high proper motion stars with observations over a long baseline (e.g. van de Kamp 1977), is removed automatically from the final RV timeseries.

- TERRA reproduces DRS RV results closely, but has been shown to outperform these particularly for M-dwarfs, where an RMS reduction of $\sim$0.3 m s^{-1} going from DRS to TERRA values was seen in several datasets (Anglada-Escudé & Butler 2012).

- In addition to RVs, TERRA outputs S-values measured from the DRS extracted spectra.

RVs obtained with TERRA have been used extensively in planet discovery papers (e.g. Tuomi et al. 2012, Anglada-Escudé et al. 2012, Arriagada et al. 2013, Anglada-Escudé et al. 2016a). Note that TERRA does not output diagnostics of the average stellar line profile, so the DRS FWHM and BIS values are still needed in the assessment of any detected RV signals (c.f. Section 4.2).

4.2 Stellar RV variability

Stellar RV jitter is an increasingly thorny issue encountered by the exoplanet community while driving down planet detection limits to ever lower RV amplitudes. This remains a major roadbloack for the detection of $\sim$10 cm s^{-1} signals caused by Earth mass planets in the habitable zones (HZ) of sun-like stars (Borgniet et al. 2015). However, with current instrumental precision at the 1 m s^{-1} level, stellar RV variability is already an important factor, and needs to be mitigated against. Keplerian RV shifts are the centre-of-mass reflex motion of the entire star. "Nuisance" signals arise from motions of (parts of) the photosphere of a star, induced by purely stellar phenomena. This RV jitter contribution is often non-random, and varies on multiple timescales (Table 4.1). In the worst case, (quasi-)periodic, purely stellar signals of this type cause a false positive planet detection. Therefore several mitigation strategies and assessments of the stellar line profiles are used to guard against this possibility (Table 4.1) and reduce the impact of stellar jitter.

A common property of the various stellar phenomena is that they distort the spectral line profiles at some level, in contrast to the parallel-shifts of the line profiles (in velocity space) due to Keplerian motion. These distortions can lead to correlations between measured RVs and diagnostics of the line profile shape, such as BIS and FWHM. Good examples of this case are given in Queloz et al. (2001), Desort et al. (2007) and Zechmeister et al. (2013). An unambiguous signature of a purely star-spot induced signal is a clear correlation between RVs and BIS. As detailed in Desort et al. (2007), detecting this signature depends on how well the stellar spectral lines are resolved. If stellar line broadening is below the instrumental resolution ($v \sin i \lesssim 3$ km s^{-1} in the HARPS case of R=115,000), correlations will typically not be detected.

In addition, or alternatively, stellar RV signals can correlate with the stellar activity index $\log(R'_{\rm HK})$ (e.g. Lovis et al. 2011). The planetary nature of RV

signals are routinely tested by searching for such correlations. In some cases photometric measurements of the star are used (e.g. Aigrain et al. 2012), although this requires simultaneous precision photometry, at a similar cadence as the RV dataset. Matches between the putative planetary RV period and periodicities in the timeseries of activity indices, BIS, FWHM & photomery are also assessed.

Sophisticated analysis techniques for dealing with stellar signals are continuously being developed to improve recovery of low-amplitude planetary signals. Fischer et al. (2016) includes an up-to-date overview of such efforts. This is an extremely active area of research, and one that has seen some controversy over the planetary versus stellar nature of prominent discoveries (e.g. Dumusque et al. 2012 versus Rajpaul et al. 2016 and Anglada-Escude et al. 2014 versus Robertson et al. 2015 versus Anglada-Escudé et al. 2016b).

For the purposes of our planet search program, we are focussed on companions with short periods (sub-day to ~ 3d, see Sections 1.2 and 1.4.3). Our cadence (~100 min) and observing run baselines (~5 days) are quite different from typical RV surveys that collect data over many years at much lower cadence. These big surveys can reduce the impact of all stellar signals in Table 4.1 with timescales less than a few years by simply averaging a large number of measured RVs. To reduce the impact of granulation, 2-3 observations spread over a night are often binned, and further binning over 2-5 nights can reduce the RV RMS from these effects to the 0.5 $m\,s^{-1}$ level (see Figure 4 in Dumusque et al. 2011a). Clearly, this is not an option for our planet search, and granulation noise is unavoidable. The work in Dumusque et al. (2011a) indicates that (for HARPS) a RV RMS of ~1.2 $m\,s^{-1}$ should nevertheless be achievable without such binning.

Table 4.1: Summary of important, non-planetary astrophysical phenomena causing stellar RV variability for main sequence FGK stars and the mitigation strategies used.

underlying cause	**effect on stellar photosphere**	**timescale**	**RV semi-amplitude** [m s^{-1}]	**mitigation strategy**	**references**
pressure waves (p-mode asteroseismic oscillations)	superposition of radial pulsation modes	few minutes	0.1 - 4 (spectral type dependent)	exposures ≥15 min	Schrijver & Zwaan (2000)
γ-Doradus (multi-periodic, non-radial, gravity mode) pulsations	complex variability of stellar disk-averaged photospheric velocity field	0.3 - 3 days	of order 1000	usually avoided as targets; testing periodicity of BIS, FHWM, photometry and their correlations with RV	Lagrange et al. (2009), Mathias et al. (2004)
variable granulation, mesogranulation, supergranulation	complex velocity evolution on different length and time scales	30 min - tens of hours	0.1 - 0.5	binning observations	Dumusque et al. (2011a), Meunier et al. (2015)
stellar magnetic activity cycle causing variations in active region coverage	variability of disk-averaged convective blueshift level	few years	0 - 10 (dep. on spectral type and mean activity level)	testing periodicity of activity indices and their correlations with RV	Lovis et al. (2011)
active regions rotating with stellar surface	suppression of convective blueshift in active regions & change in flux-balance between red- and blue-shifted parts of photosphere	stellar rotation period (tens of days)	0.2 - 3 for low activity stars; up to 100 - 200 for highly active, rapid rotators	testing periodicity of BIS, FHWM, photometry and their correlations with RV	Dumusque et al. (2011b), Desort et al. (2007)

Other stellar signals are expected to be less problematic for our programme:

- We can mitigate against stellar p-mode noise by using 15 minute exposures.
- Stellar activity cycle variations are irrelevant for observing runs of a few days.
- The vast majority of our targets are not young, fast rotators. Their stellar rotation periods should be on the order of tens of days; well separated from the short-period planetary signals of interest. Over ~5 nights, low-amplitude linear or quadratic RV trends may arise from stellar rotation.
- None of our targets are known γ-Doradus pulsators. Only the earliest spectral types (~ F4-F8) have the potential to show such pulsations; I discuss this in detail in Section 5.3.4.

There is evidence that TERRA RVs can be somewhat less affected by purely stellar signals than CCF derived values (Anglada-Escudé & Butler 2012), a point that will be confirmed for our programme in Chapter 5.2.2.

4.3 Instruments used in our RV programme

A priori it was unknown whether our planet search would yield large RV amplitude, HJ-type detections, or much lower amplitude signals of low-mass, close-in planets. Archival RV data for a subset of targets strongly disfavored the former scenario, while the latter would arguably be more scientifically interesting (Section 1.5). As discussed in Section 4.4, RV precision on the order of 1-2 $\mathrm{m\,s^{-1}}$ enables the detection of close-in planets down to a few Earth masses in a small number of nights. There are limited choices for astronomers (particularly European teams) wishing to conduct a RV planet search at this

precision. HARPS at La Silla (Sect. 4.3.1) is the only suitable Southern Hemisphere instrument that we had access to. In the Northern Hemisphere the OPTICON scheme[2] provides access to the SOPHIE (Sect. 4.3.2) and HARPS-N instruments, with 2 and 1 $m\,s^{-1}$ precision respectively. As described in Section 4.4, our simulations predicted a relatively small impact of this change in precision for planet detection limits. We typically applied for telescope time on both facilities. HARPS-N has a significantly higher oversubscription rate than SOPHIE, therefore we were awarded SOPHIE time for almost all our Northern Hemisphere observations.

4.3.1 The HARPS and HARPS-N spectrographs

HARPS

The **H**igh **A**ccuracy **R**adial velocity **P**lanet **S**earcher (Mayor et al. 2003) has been attached by fibre-link to the Cassegrain focus of the 3.6 meter telescope at La Silla since 2003. HARPS is a cross-dispersed echelle spectrograph with 2 CCD detectors covering 72 spectral orders from 3780 to 6910 Å. In its High Accuracy Mode (HAM), the resolving power is 115,000. A lower-resolution, higher efficiency "EGGS" mode is not used for precision RV measurements. The 2 fibres feeding the spectrograph have on-sky diameters of 1'', and can be used in *ThoSimult* or *objAB* mode. HARPS was carefully designed for exquisite instrumental stability and RV efficiency to ensure low-mass planet detection capability. A key design feature is that the entire spectrograph is sealed in a thermally isolated vacuum vessel, ensuring long term, extreme levels of mechanical, temperature and pressure stability. On the timescale of one night

[2]http://www.astro-opticon.org/

both the total instrumental drift and the RV stability measured for stellar targets are well below 1 $m\,s^{-1}$; the precision can be as good as 0.2 $m\,s^{-1}$ on such short timescales (Mayor et al. 2003). Crucially, the long-term stability of HARPS on timescales of several years is known to be better than 1 $m\,s^{-1}$, based on observations of RV standard stars. A 6-year timeseries of HD10700 showed RV scatter of only 0.9 $m\,s^{-1}$ (Pepe et al. 2011). The instrumental stability must be below this level. The exact value is unsure due to the unknown stellar variability contribution that is present even for such extraordinarily stable stars.

Detailed descriptions of the instrument can be found in Pepe et al. (2002a) and Pepe et al. (2003). Its installation (Mayor et al. 2003) marked a milestone in the history of precision RV searches: the key breakthrough of long-term $< 1\ m\,s^{-1}$ stability was combined with an instrument primarily dedicated to RV planet searches, fed by a large telescope with excellent observing site conditions. This has enabled a plethora of planet discoveries in the Southern Hemisphere, including many super-Earth and Neptune-mass planets. Many of the HARPS surveys and results are described in the ongoing publication series "The HARPS Search for Southern Extra-solar Planets" (currently comprising 40 papers), but this is by no means a complete listing of all discoveries enabled by the instrument. Particular highlights in the HARPS planet haul to date have been: the first observed multiple planet system composed only of Neptune-mass objects (Lovis et al. 2006); a system with 3 super-Earths (Mayor et al. 2009); measurements of the mass of super-Earths detected in transit - establishing their density was consistent with rocky composition (Queloz et al. 2009); and (super-)Earth mass planets in the habitable zones of M-dwarfs (e.g. Anglada-Escudé et al. 2013), most spectacularly for Proxima Centauri (Anglada-Escudé et al. 2016a).

HARPS was upgraded in June 2015, resuming regular operations on 2015-06-03 (Lo Curto et al. 2015). The most important change was the installation of

a new octagonal fibre link. This followed positive results of a similar upgrade to SOPHIE (see Sect. 4.3.2), demonstrating that octagonal fibres provide improved scrambling compared to fibres with circular cross-section. Results from HARPS (Lo Curto et al. 2015) confirmed this, demonstrating that the instrumental profile became less sensitive to de-centering of target stars on the fibre from guiding or focus offsets. The instrument throughput was also increased significantly. A downside of the upgrade was a one-off instrumental profile change, leading to an RV offset of up to 20 $\mathrm{m\,s^{-1}}$ in the timeseries of all stars observed before and after June 2015 (Lo Curto et al. 2015). Unfortunately the offset value varies significantly with stellar spectral type and FWHM. Therefore the HARPS team recommends treating pre- and post-upgrade measurements as independent datasets, and adding an RV offset as a free parameter in the analysis. The observations of our programme were all carried out after this upgrade. In several cases, we combined our data with pre-upgrade archival RVs, for which this effect had to be accounted for.

HARPS-N

The HARPS-N spectrograph was conceived and designed as a Northern Hemisphere "clone" of HARPS. A description of the instrument can be found in Cosentino et al. (2012). All basic properties are as outlined for HARPS, including wavelength coverage, resolution, simultaneous reference method, octagonal fibres for improved scrambling, and RV stability. In our RV programme, we only acquired a limited amount of HARPS-N data for a single target. HARPS-N was installed at the 3.6-m Telescopio Nazionale Galileo (TNG) at the Roque de los Muchachos Observatory in 2012 (Cosentino et al. 2012). It has been of great value for mass measurements of transiting planets detected by the Kepler mission (e.g. Dumusque et al. 2014) and for RV discovery of planets in the

Northern Hemisphere (e.g. Niedzielski et al. 2015). High profile results include the study of the USP planet Kepler-78 (Pepe et al. 2013), which confirmed an Earth-like density for this 1.9 $M_\oplus$, 1.2 $R_\oplus$ planet, and the discovery of a transiting multi-planet system around the very close-by (6.5 pc) star HD219134 (Motalebi et al. 2015).

Observing strategy for HARPS

To prevent moonlight contamination, we avoided bright-time observations whenever the lunar RV was within 2 FWHM of a stellar systemic velocity. We used the High Accuracy Mode with default detector readout settings, and carried out the standard set of daily afternoon calibrations (Bias, Flats, ThAr exposures) as required for HARPS high precision RV measurements.

In contrast to SOPHIE (Sect. 4.3.2), HARPS is not set up to automatically read out exposures once a specified SNR level is reached. I will use "SNR" as shorthand for the signal-to-noise per wavelength pixel of the extracted spectrum at 5500 Å throughout the rest of this thesis, unless noted otherwise. This is the conventional signal-to-noise metric used throughout the RV exoplanet literature. To average asteroseismic oscillations, we consistently observed with a minimum exposure time of 15 minutes. In typical conditions at La Silla (1" seeing and no cloud cover), SNR $\geq$ 100 is achieved in 15 minutes even for our faintest Southern Hemisphere targets (V $\sim$ 9.5). This corresponds to photon-limited internal RV uncertainties $\leq$ 1 m s^{-1}. We increased exposure times to maintain precision as close to 1 m s^{-1} as possible during poorer observing conditions encountered in some of our runs. A maximum exposure time of 40 minutes was imposed as an additional constraint. This is because systematic effects arising from very long exposures outweigh the advantage of improved photon noise:

- For the shortest periods of interest in our USP search (~ 4 hours) the orbital reflex RV would be "smeared out" in exposures that span significantly longer than 40 minutes. This threatens the detectability of particularly interesting planetary signals, and should therefore be avoided.

- The RV smearing from Earth's rotation also becomes significant for long exposures. The magnitude of the effect depends on the details of a given observation (i.e. stellar and observatory coordinates; observing epoch). For reference, the BERV during our HARPS observations changed by up to ~ 80 m s^{-1} over 40 min. This effect leads to stellar line broadening and therefore reduced RV precision.

The exact value for the maximum exposure length is somewhat arbitrary. In practise, exposure times were $\lesssim 20$ min during the vast majority of our HARPS observations, as observing conditions ranged between adequate and ideal. The maximum exposure limit is more relevant for SOPHIE observations (Sect. 4.3.2).

HARPS CTI effect

CCD detectors are read out by transfering charge from pixel to pixel. The process is never perfect, with very small fractions of the charge becoming trapped before readout. The so-called charge transfer inefficiency (CTI) quantifies this charge trapping and is a well-studied property of CCDs (e.g. Dolphin 2000). CTI can be a significant problem for precision RVs, for which line centroid shifts corresponding to ~ 0.001 pixels are measured. Unless the CTI is negligible, purely instrumental systematic changes to the stellar line profiles are induced (Bouchy et al. 2009a). The magnitude of the effect correlates with signal strength on the CCD, which causes a spurious dependence of the measured RV on the SNR. For the SOPHIE CCD the CTI effect is substantial: it reaches a

peak-to-peak amplitude of 20 m s^{-1} over the SNR range covered by typical observations (Bouchy et al. 2009a). For the HARPS CCD a much smaller effect on the RVs is seen. The HARPS-N DRS automatically corrects for its low-level CTI effect (Cosentino et al. 2014), but for HARPS this is not implemented (Lovis, private communication). Berdinas et al. (2017) used high-cadence observations of a single M-dwarf RV standard star to fit an empirical correction as function of SNR at echelle order number 60 (~6120 Å), SN_{60}. Note that this was chosen as a more appropriate metric for M-dwarfs compared to the SNR at 5500 Å, because their SEDs peak in the red compared to FGK stars. The empirical relation derived in Berdinas et al. (2017) between the CTI correction, $\Delta RV_{\rm CTI}$, and SN_{60} has the form

$$\frac{\Delta RV_{\rm CTI}}{[\rm m\ s^{-1}]} = 4.92\ -\ 1.31\ \ln(SN_{60}). \tag{4.1}$$

Note the constant term is irrelevant for correcting differential RVs of a given target, but is needed to provide an adequate fit of the empirical correction itself. This correction removes the spurious red-shift of HARPS RVs at low signal to noise levels. The TERRA package provides RV timeseries with the correction as an additional output. Its use has been limited to M-dwarf stars as the validity for other spectral types has not been tested (Anglada-Escude, private communication). To my knowledge, there has not been a publication to date that addresses this. Equation 4.1 shows that in a worst-case scenario, with SN_{60} ranging from 50 to 200 due to e.g. highly variable observing conditions and exposure lengths, the differential correction is of order 2 m s^{-1}. For more typical cases with relatively consistent signal to noise levels, the corrections will be below 1 m s^{-1}. The effect is therefore of low importance, in particular for observing programmes where SNR varies quasi-randomly. In the case of our high cadence observations and those of Berdinas et al. (2017), this is not necessarily the case: changing observing conditions can cause SNR trends over

a given night, and significant jumps in the mean SNR achieved from night to night. Therefore the CTI effect introduces low-level red noise at the 1 $m\,s^{-1}$ level, which could interfer with the detection of our lowest RV amplitudes of interest (Sect. 4.4). Therefore I extended the empirical CTI correction to FGK spectral types, as needed for our observing programme.

Based on the data compiled in Soubiran et al. (2015) I selected suitable RV standards: FGK stars with a large number ($\gtrsim 400$) of archival, publicly available HARPS observations, with SN_{60} spanning as large a range as possible. Four stars of spectral type F9 to K5 matched these criteria: HD1581, HD85512, HD154577 & HD190248. I used observations taken before the HARPS instrumental upgrade to avoid the corresponding RV offset, and limited the SN_{60} range from 30 - 500. I normalised the TERRA RVs of these datasets to their median values at $100 < SN_{60} < 130$. For a consistent comparison to the work by Berdinas et al. (2017), I kept the functional form of Equation 4.1. I used non-linear least-squares fitting with the Levenberg-Marquardt method, which resulted in

$$\frac{\Delta RV_{\rm CTI}}{[\rm m\ s^{-1}]} = 4.88\ -\ 1.03\ \ln(SN_{60}). \tag{4.2}$$

I also tested fitting each RV standard separately. The resulting correction laws agreed within error, and no clear dependence on stellar spectral type was seen. As shown by Figure 4.3, this correction is very similar to that in Berdinas et al. (2017): the differences are below 0.5 $m\,s^{-1}$ even at the lowest and highest signal to noise levels. The HARPS datasets of our programme had typical SN_{60} ranges of ~50-120, when combining data from multiple observing runs with highly variable conditions. From Equation 4.2, this corresponds to differential corrections up to 0.9 $m\,s^{-1}$. However, the photon-noise limited RV precision at the low SNR end was at best 2 $m\,s^{-1}$. In practise, I found the CTI correction was not an important effect in our HARPS observations, as illustrated by a typical example shown in Figure 4.4.

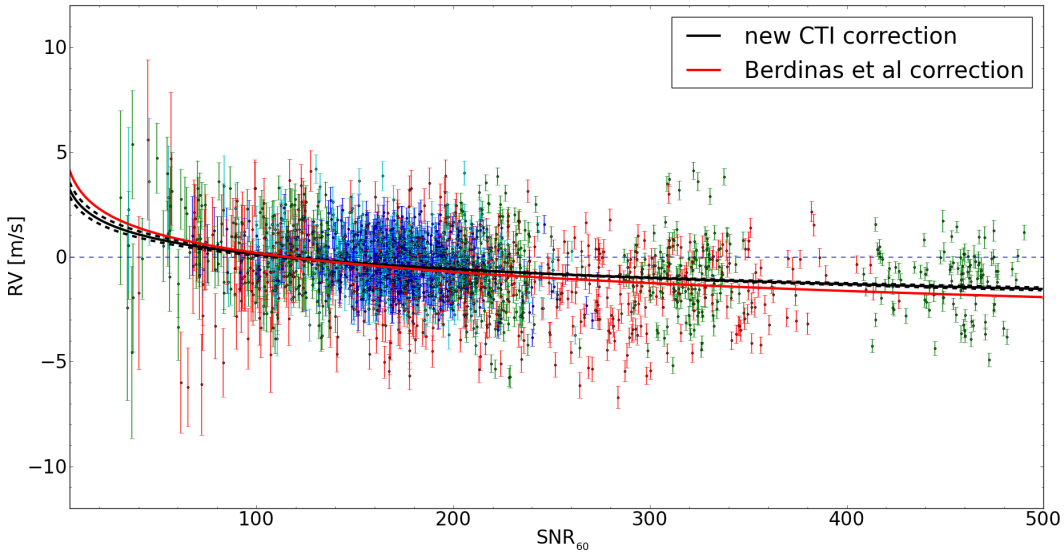

Figure 4.3: Best fit of my CTI correction to archival FGK RV standard datasets (colour-coded), compared to the Berdinas et al. (2017) correction derived from observations of a single M-dwarf. To allow a like-for-like comparison, the latter correction law has been offset by a constant so that it passes through $\Delta RV_{CTI} = 0$ at $SN_{60} = 115$, matching the normalisation I applied to the RV data.

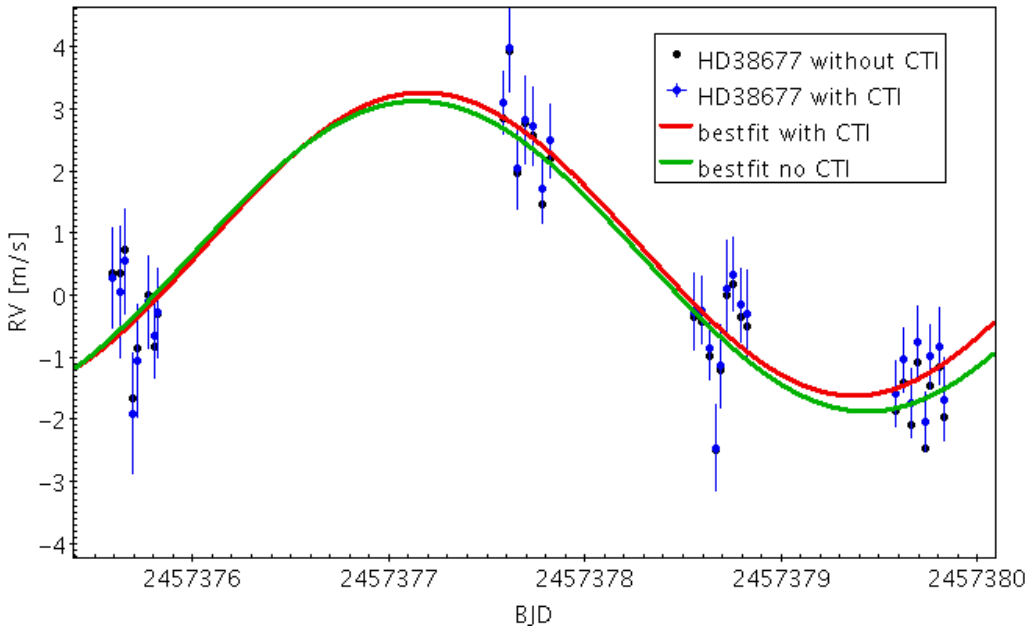

Figure 4.4: Example of the small effect of the CTI correction on a HARPS dataset acquired during our programme. Changes in observing conditions between nights led to a change in mean SNR, and therefore an RV offset of 0.5 m s^{-1} during the final night of the run. For comparison, the resulting change in the best fit planetary signal is also shown. See Section 5.4 for details on the analysis and fitting of this signal.

4.3.2 The SOPHIE spectrograph

The SOPHIE spectrograph (Perruchot et al. 2008) began operations in October 2006 at the 1.93-m telescope of the Observatoire de Haute-Provence (OHP). The first detection of an exoplanet around a sun-like star (51 Peg) was achieved with this telescope, using SOPHIE's predecessor ELODIE (Mayor & Queloz 1995). SOPHIE is a fibre-fed cross-dispersed, environmentally stabilized echelle spectrograph covering the wavelength range from 3872 to 6943 Å, with spectral resolving power of 75000 in high-resolution (HR) mode and 39000 in high-efficiency (HE) mode. A detailed description of the instrument is given in Perruchot et al. (2008). As for HARPS, two optical fibres are used allowing *ThoSimult* or *objAB* mode observations. The fibres have a sky acceptance of 3″, matching the much poorer median seeing conditions at OHP compared to La Silla. A near-identical version of the DRS available for HARPS(-N) provides automated data reduction, with analogous data products.

Although OHP does not experience weather and seeing conditions on par with the TNG and La Silla sites, SOPHIE is a work horse facility in the RV search for Northern Hemisphere exoplanets (Bouchy et al. 2009c). Follow-up and RV validation of transit detections from the SuperWASP, HATNET and Kepler surveys are carried out routinely (e.g. Pollacco et al. 2008, Boisse et al. 2013, Santerne et al. 2012). SOPHIE has also been used to measure planetary system obliquities (e.g. Bourrier et al. 2015). In the first years of operation RV precision was limited to 5 - 6 m s^{-1} at best (e.g. Boisse et al. 2009). This was primarily caused by insufficient scrambling within the optical fibre link and a high sensitivity to illumination variations (Boisse et al. 2010), leading to systematic effects from seeing variations, guiding decentering, telescope defocusing and atmospheric dispersion. These issues were mitigated by the insertion of a fibre link with octagonal cross-section and greatly improved scrambling properties (Bouchy et al. 2012) on the 17th of June 2011. Subsequent data is sometimes

referred to as SOPHIE+ in the literature, with demonstrated RV precision of 1 - 2 $m\,s^{-1}$ in HR mode on time scales of tens of days for several RV standard stars (Bouchy et al. 2012). The HE mode offers high throughput but precision remains limited to 3 - 4 $m\,s^{-1}$ (Bouchy et al. 2012), primarily due to the absence of a double-fibre scrambler. Typically this mode is used for low precision observations of stars with $V \gtrsim 12$.

Observing strategy for SOPHIE

We followed the standard daily calibration plan (Bias, Flats, ThAr exposures) used for all SOPHIE RV observations. We observed in HR mode, since our Northern Hemisphere targets of interest were $V \lesssim 10$. *objAB* mode was used as the default, with Fibre A collecting starlight and Fibre B recording the sky spectrum at a separation of 2′. In contrast to simultaneous ThAr injection into the second fibre, *objAB* mode allows reliable monitoring of realtime (cumulative) counts, via SOPHIE's flux meter. This is set up to automatically read out exposures once the desired SNR level is reached, i.e. the exposure length is adapted in realtime to compensate for extinction and seeing losses. We imposed a maximum exposure time of 40 minutes, as for HARPS. With SOPHIE, an additional systematic effect for long exposures becomes significant: the instrumental drift is much larger than for HARPS (Figure 4.5). In *objAB* mode it is only tracked before and after science exposures. Therefore the precision of drift corrections will have an inverse relationship to the length of science observations.

With the *objAB* setup, the sky spectrum from Fibre B can be used to monitor moonlight contamination levels, and there is no contamination from saturated ThAr lines leaking into stellar spectrum on the CCD. A more consistent SNR level can be achieved for a given target throughout the run using the realtime exposure time adjustment, even in the relatively variable observing conditions

at OHP. This reduces systematic effects from imperfect CTI corrections, and prevents read-out if flux levels are too low for a useful measurement. Both of these issues and the ThAr contamination effect are potentially much greater concerns than for HARPS (Section 4.3.2), due to the typically lower SNR achieved with SOPHIE, combined with worse and more variable OHP observing conditions.

As a compromise between photon noise level and sufficiently high observing cadence, we aimed for SNR = 60 for our faintest target (V = 9.6), and SNR = 190 for the brightest star (V = 5.7), leading to RV photon noise contributions of 3 - 1 m s^{-1}. This is sufficient given SOPHIE's instrumental stability of 2 m s^{-1} and since coincidentally RV precision was limited by stellar rotation broadening for the fainter targets. Under excellent observing conditions, the desired SNR levels were reached on timescales of a few minutes for our brighter targets. When this occurred, consecutive short exposures of these targets were taken in blocks of $\sim$ 15 min length. Averaging over stellar p-mode noise could then be achieved by binning the corresponding datapoints.

SOPHIE instrumental drift

In the absence of simultaneous ThAr injection into Fibre B, the instrumental drift was tracked by frequent calibration ThAr exposures, taken after each stellar observing block (i.e. at least 2 per hour). This is more than sufficient for typical SOPHIE drifts of $\leq$ 3 m s^{-1} h^{-1} (Bouchy et al. 2009c) and allowed detection of any anomalously large drifts on short timescales. We calibrated during telescope slews between targets, avoiding a significant increase of overheads. During one night of our programme (2015-04-29) an anomalously large drift (up to 13 m s^{-1}h^{-1}) was observed (Figure 4.5). This was probably related to a replacement of the air conditioning unit in the instrument electronics room and associated temperature variations. Due to frequent drift measurements this

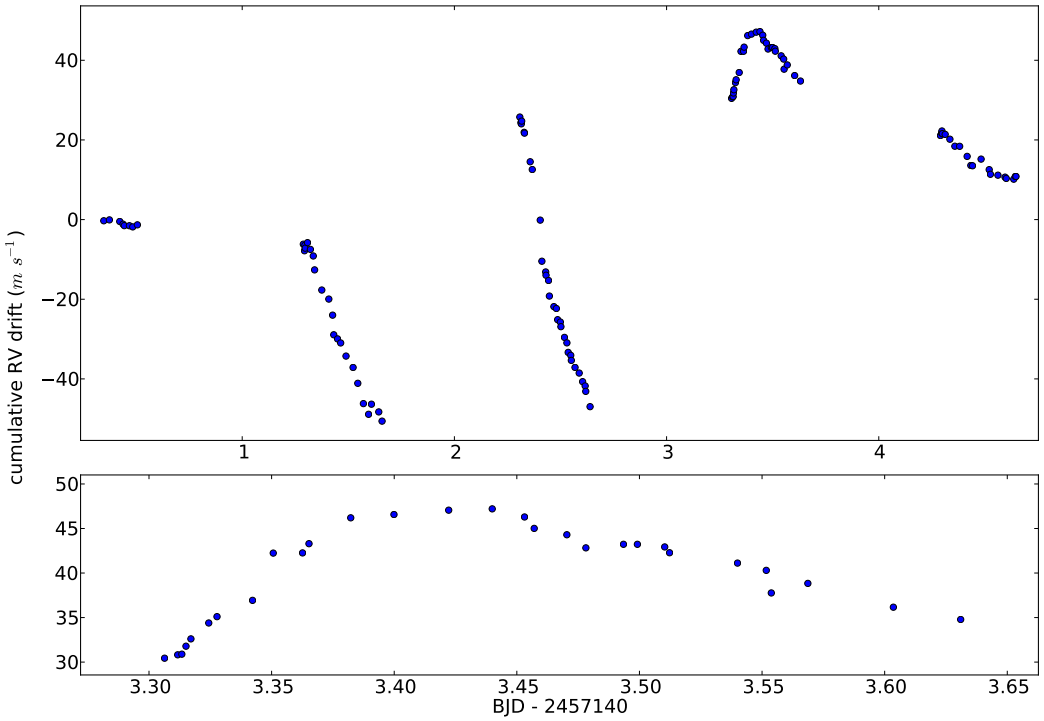

Figure 4.5: Time series of SOPHIE's instrumental RV drift throughout our first 2015 run (top), and during one night (2015-04-30, bottom). Data gaps and sparsely sampled regions correspond to interruptions of observations by daylight and bad weather. The photon noise uncertainties on the drift measurements are at the 0.07 m s^{-1} level.

did not have a significant effect on RV precision, as illustrated by Figure 4.5. For *objAB* observations the DRS pipeline simply applies the instrumental drift value from the preceding ThAr calibration. I followed the standard practice (e.g. Boisse et al. 2010) of improving the drift corrections by applying linear interpolation of the measured ThAr drifts, evaluated at the science exposure midpoints. For several epochs this resulted in non-negligible RV corrections of up to 4 m s^{-1}. It is worth noting that the instrumental drift of SOPHIE is much greater compared to HARPS, primarily because the entire instrument is not encapsulated in a thermally isolated vacuum vessel.

SOPHIE CTI effect

In the specific case of the SOPHIE CCD, charge transfer inefficiency results in a high amplitude spurious blue-shift for exposures with low SNR. The magnitude of the effect is much greater than for HARPS, reaching $\sim 10\ \mathrm{m\,s^{-1}}$ at SNR $\sim$ 70 and $\sim 20\ \mathrm{m\,s^{-1}}$ at SNR $\sim$ 40. For a detailed discussion, see Bouchy et al. (2009a). In spite of using the SOPHIE flux-meter as described in Section 4.3.2, we could not achieve perfectly consistent SNR for several targets throughout our observations, due to highly variable conditions at OHP and our maximum exposure time limit of 40 minutes. Therefore a correction of the CTI effect was vital to avoid systematic RV shifts in excess of both the signals of interest and the intrinsic instrumental stability. Unfortunately, automated correction of the CTI by the SOPHIE DRS is not implemented (Bouchy, private communication). In the literature the empirical correction detailed in Santerne et al. (2012) is sometimes used. However, this was derived from HE mode observations and has a quoted range of validity of SNR = 10-50. Therefore a new empirical correction law was required for our HR mode, SNR $\sim$ 40-200 observations. Stable RV standard stars observed over a wide range of SNR are needed to fit a relation between the CTI correction, ΔRV_{CTI}, and SNR. I followed the methodology of Santerne et al. (2012) to derive an updated power-law correction, as I found the functional dependence of ΔRV_{CTI} on SNR to be broadly consistent between HR and HE data.

I used the publicly available HR mode datasets of HD 221354, HD 30708, HD 185144, HD 139324, HD 109358 & HD 5372; observed after the SOPHIE+ upgrade with SNR = 17 - 260 and identified as RV standards via Bouchy et al. (2012). After normalising all datasets by their median RV at SNR > 140, a power law was fit using non-linear least squares (Fig. 4.6). The best fit relation is

$$\frac{\Delta RV_{\mathrm{CTI}}}{[\mathrm{m\ s^{-1}}]} = -3170 \quad SNR^{-1.37} \ + \ 2. \tag{4.3}$$

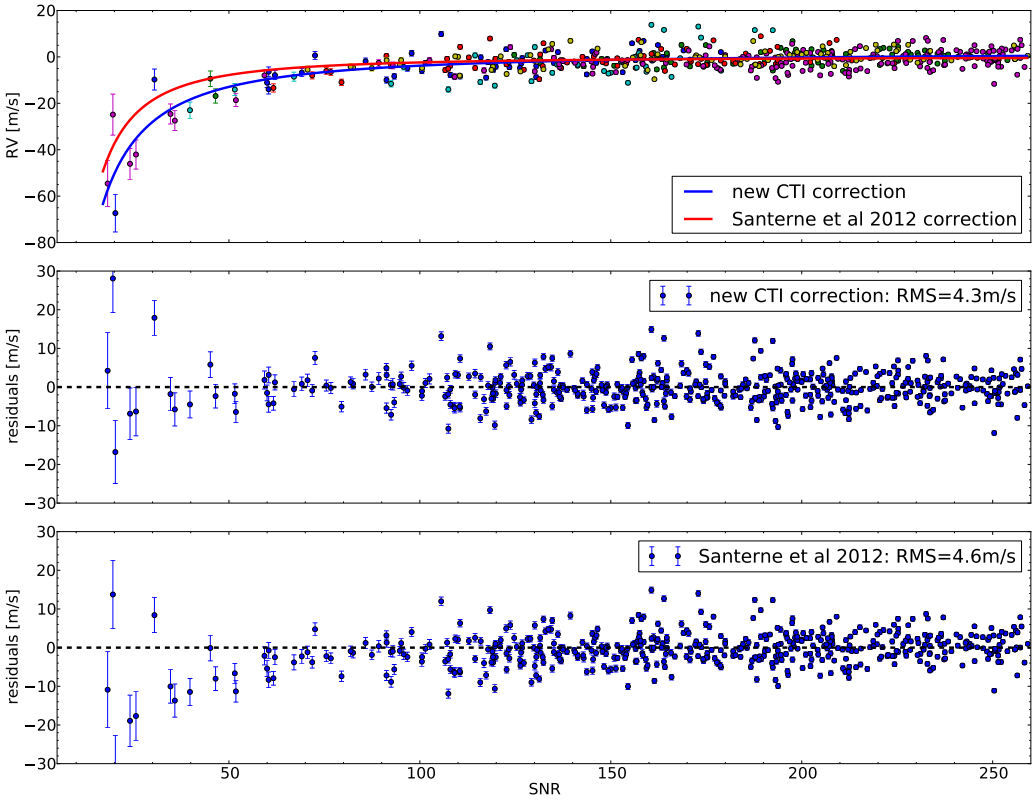

Figure 4.6: Best fit and residuals of my CTI correction to archival RV standard datasets compared to the Santerne et al. (2012) correction. Different RV standard stars are colour-coded in the top panel.

The addition of a constant term removes a spurious trend in the RV standard observations' residuals: the effect does not asymptotically tend to zero at high SNR observations when normalising to SNR = 140-260. When applying the correction to our own observations, this constant is clearly irrelevant, as only differential RVs are required. As shown in Figure 4.6, the updated relation is a significantly better fit to the HR data in the critical low SNR regime than the Santerne et al. (2012) relation. At SNR $\lesssim$ 60 there is a dearth of RV standard data. Additional observations in this regime are advisable to further improve the CTI correction. Applying Equation 4.3 to our targets' RV timeseries led to differential RV corrections of up to 10 m s^{-1}.

Flux-weighted exposure midpoints

SOPHIE's exposure meter provides the flux-weighted midpoints of all observations, storing these midpoints in the headers of the DRS output files. Unlike HARPS, the current SOPHIE DRS pipeline does not use this information further: the BERV correction is not updated accordingly (Bouchy, private communication). In good observing conditions this can be tolerated in the context of SOPHIE instrumental precision of 2 $\mathrm{m\,s^{-1}}$. Observations taken in poor conditions are often simply discarded if large datasets are available (e.g. Borgniet et al. 2014). During our program we aimed to retain such observations if their SNR was sufficient; avoiding excessive weather losses. For some of our observations affected by highly variable clouds and/or seeing, mid-point shifts on the order of ~200 seconds occurred, translating to a change of the BERV of $\sim 4\ \mathrm{m\,s^{-1}}$. I therefore calculated the BERV at the fluxweighted midpoints via the *PyAstronomy* implementation of the *helcorr* routine[3]. Then I corrected the measured RVs by the difference between the updated BERV and the "original" BERV at the exposure half-way points. It should be noted that *ThoSimult* SOPHIE exposures cannot be fully corrected in this way. In this mode, the exposure meter measures flux from both the target star and the ThAr lamp. Therefore any midpoint shifts of the stellar exposure are diluted. This is another reason to use *objAB* mode for SOPHIE observations.

Moonlight contamination

In contrast to HARPS, one of our SOPHIE runs (2015A) was carried out during bright time, while at the same time the velocity difference between some of our targets and the moon was small enough (within 1 FWHM) to be of concern. Background levels were boosted during this run by thin cloud cover scattering moonlight. We recorded significant background levels for several exposures,

[3]http://www.hs.uni-hamburg.de/DE/Ins/Per/Czesla/PyA/PyA/pyaslDoc/aslDoc/baryvel.html

i.e. a clear CCF detection of the lunar RV from Fibre B, close to the stellar RV from Fibre A. The targets affected were HD120666 and HD184960.

The SOPHIE DRS is set up to correct for this contamination effect. Since Fibre B records the sky spectrum in *objAB* mode, the DRS pipeline extracts and reduces this in the same way as for Fibre A, including CCF computation. This sky CCF can be subtracted from the stellar CCF, and the final RV is measured from this corrected CCF. The entire procedure is automated by the SOPHIE DRS, and has been implemented repeatedly in the literature (e.g. Barge et al. 2008, Pollacco et al. 2008). Figure 4.7 demonstrates the importance of this step and the level of contamination for HD120666 observations. I found that without the correction the CCF profiles of HD120666 were heavily distorted, with spurious RV shifts of up to 90 $\mathrm{m\,s^{-1}}$. For HD184960, the effect is more marginal: a significant sky signal was only detected in 3 exposures, according to the criteria of Díaz et al. (2012): Fibre B CCF contrast > 1% and sky spectrum SNR > 3. For these points the difference between original and corrected RVs ranged from 6-17 $\mathrm{m\,s^{-1}}$. Note that TERRA RVs currently cannot be corrected for moonlight contamination reliably (Anglada-Escude, private communication). For HD120666 only DRS RVs were therefore used. For HD184960 I simply discarded the 3 contaminated measurements to have a consistent comparison between the DRS and TERRA timeseries.

Summary of SOPHIE RV corrections

The contrast between SOPHIE and HARPS in terms of necessary corrections is clear. Typical differential RV corrections for SOPHIE targets can be summarised as follows, based on our experience from the 2015A and 2015B runs. 2015A suffered from very poor observing conditions with highly variable cloud cover and seeing, while the 2015B run was mostly cloud-free with consistently good seeing.

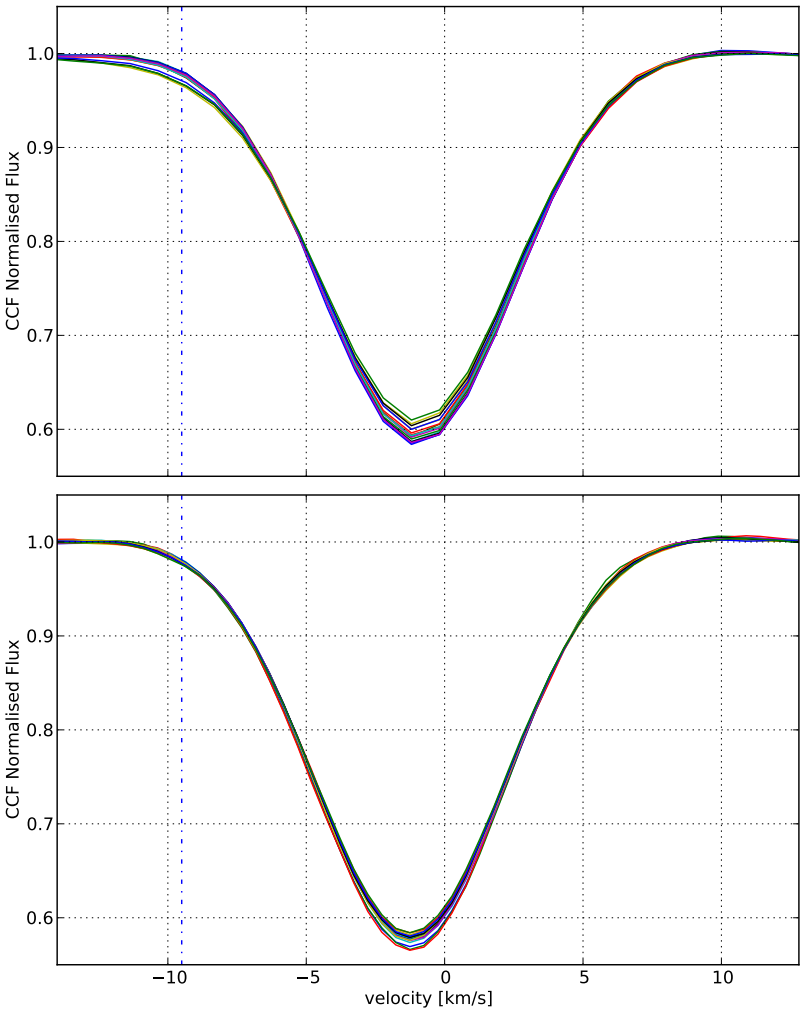

Figure 4.7: HD 120666 CCFs for all 2015A exposures, before (top) and after (bottom) the correction of sky contamination. Sky background levels varied throughout the run, and contamination peaked at the lunar radial velocity of ~ -9 km s^{-1} (dashed).

- **Moonlight contamination corrections** were important for only 1 target, with up to 90 m s^{-1} shifts.
- **Instrumental drift corrections** were highly variable, ranging between 0 and 11 m s^{-1}.
- **Flux-weighted exposure midpoint corrections** varied between 0 and 10 m s^{-1}, but were typically at the 1 m s^{-1} level during the 2015B good weather run.

- **CTI corrections** varied between 0 and 8 m s^{-1} during 2015A, and between 0 and 3 m s^{-1} in 2015B, reflecting the reduced SNR variability and higher overall SNR levels achieved in better weather conditions.

As a "worst case" example, the RMS of the HD120666 dataset in 2015A was reduced from 34.8 to 6.6 m s^{-1} after I had applied all corrections. In this case the moonlight contamination correction alone reduced the RMS by 25 m s^{-1}. A more typical example in the 2015B run is the dataset of HD11130, for which the RMS was reduced from 3.4 to 2.8 m s^{-1} after all corrections.

4.3.3 The chromatic systematic effect

Bourrier & Hébrard (2014) and Berdinas et al. (2016), hereafter B16, identified a significant intra-night systematic effect on observations made with the HARPS-N spectrograph. B16 provides an in-depth assessment of this, using high cadence data from a wide binary M-dwarf system. The root cause appears to be an incomplete correction of differential atmospheric refraction by the atmospheric dispersion corrector (ADC), causing colour-dependent (chromatic) flux-losses. The chromatic distribution of stellar light entering the fibre is therefore time-dependent.

I followed the methodology of B16 to quantify the effect for our observations. In brief, I measured the "spectroscopic chromatic index" κ, defined as the slope of a linear portion of the pseudo spectral energy distribution (pSED) about a specified wavelength (see Fig. 4.8). κ values are given in counts per micron, hereafter counts μ^{-1}. The pSED is obtained from high resolution spectra by computing the sum of counts in each spectral order, normalised by the counts in a reference order R (Fig. 4.8). I measured κ as described in B16, with modifications to account for the spectral types of our targets (FGK instead of the M-dwarfs in B16). For SOPHIE data, I normalised the pSEDs to the mean flux in echelle order 20, and the linear slope was fit over orders 15 to 25 (Fig. 4.8).

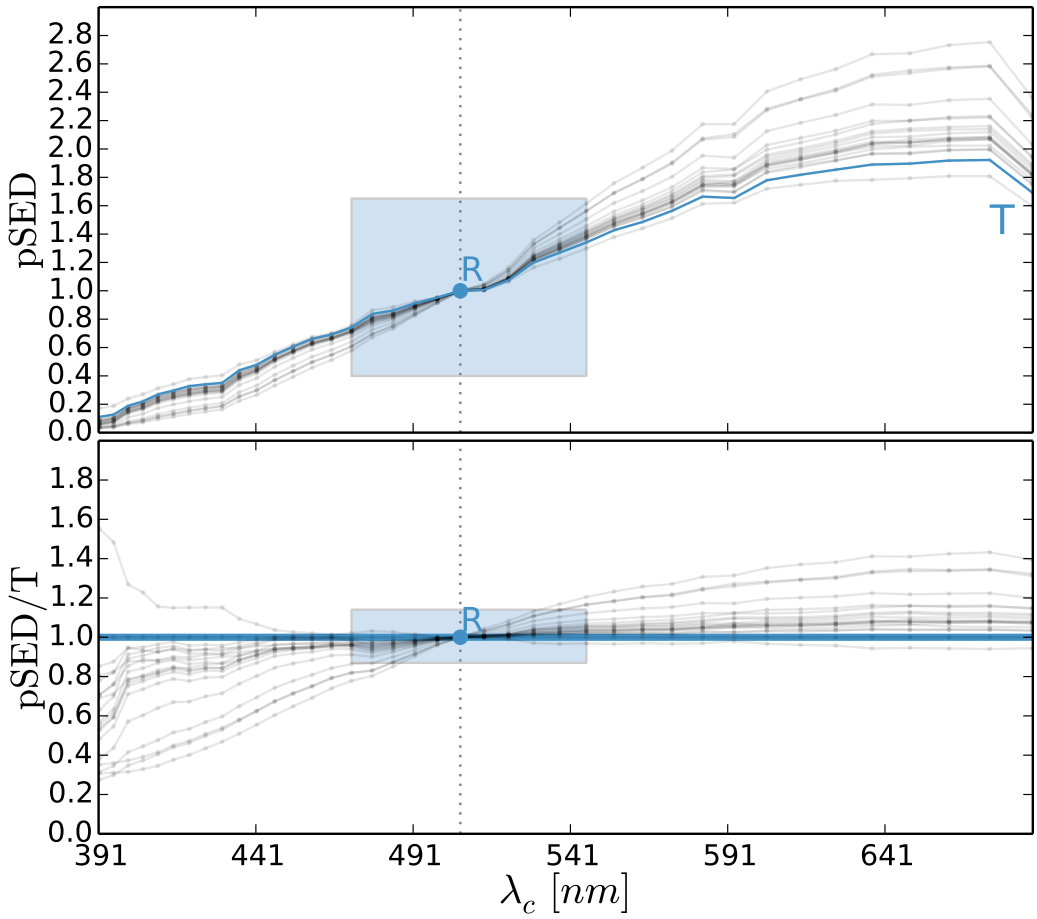

Figure 4.8: Top: observed pSEDs of HD 121131 SOPHIE observations, normalised to reference order R. Each echelle order is represented by a dot. Colour-dependent flux losses are obvious, and are seen in the SEDs of all targets. Bottom: normalised SEDs divided by the first observation (T). κ was fit in the approximately linear, highlighted region.

For HARPS, I used a linear fit over orders 54 to 64, with order 59 as the pSED normalisation reference. A variable pSED slope does not affect the intrinsic stellar line profiles, but can cause an apparent change during CCF computation. This is because CCFs of the individual spectral orders are weighted by their flux levels in the computation of the final average CCF. Each order's CCF has a subtly different profile. Variable weighting arising from the chromatic effect will then cause a variable average CCF for each star (B16). The CCF's FWHM is most strongly affected, but a non-negligible, spurious RV shift is also induced.

Following the assessment of this effect in Bourrier & Hébrard (2014), the DRS version installed at HARPS and HARPS-N now corrects for it (Cosentino et al. 2014). This is done automatically via continuum flux normalisation relative to templates created from archival observations of different spectral types. The step is done before the orders' CCFs are weighted and combined. However, the correction is not done for M-dwarfs, due to "division by zero" problems at the bluest wavelengths with very little stellar flux. **This DRS correction is not implemented for any spectral type in the SOPHIE DRS.**

TERRA RVs are obtained via a simultaneous fit of Doppler shift and parameters describing the continuum (flux normalisation term, see Section 4.1.3). They are therefore not sensitive to the issue described above, as shown by the tests done in B16. TERRA data from one of the observing runs in B16 nevertheless showed a strong RV-κ correlation. Intriguingly this behavior was not seen in a repeated observing run, including similar observations of the same target. The cause for this remains unknown, especially since B16 found no evidence that the intrinsic HARPS-N instrumental line profile is changed by the chromatic effect. The efficient scrambling in the fibre-link of the instrument prevents this.

I will show in Chapter 5 that the chromatic effect has a strong impact on SOPHIE CCF results. As for HARPS-N, there is a strong correlation between FWHM and κ. Crucially, the measured RVs are also affected. Given the scheduling constraints for our programmes, we cannot avoid observing over a relatively large range of airmass (i.e. atmospheric dispersion). Therefore this systematic effect is an important consideration in our SOPHIE analysis. In contrast, I will show that our HARPS data (TERRA and DRS RVs) are not affected, in accordance with the results of a separate investigation (Anglada-Escude, private communication).

4.4 Achievable detection limits

Clearly the detection limits for a planet search are driven by the RV precision per measurement, the cadence of observations, and the total number of observations per target. The landscape of possible scheduling permutations of any RV programme is vast, but practical constraints limit this somewhat. To make full use of the 1-2 m s^{-1} instrumental precision of our facilities, the photon-noise contribution to RV uncertainty should be at the ≤ 1 m s^{-1} level, corresponding to SNR ≥ 100. Depending on our targets' brightness, observing conditions and which facility is used, this is achieved within 5 - 30 minutes and within 10 - 20 minutes for the majority of cases. As discussed in Section 4.2, a minimum "block" length of 15 min per target is required to average over asteroseismic oscillations. We found that the total telescope overhead time between observing blocks with different targets is on the order of 5 minutes for both SOPHIE and HARPS. Therefore a typical observation including overhead will last 20 min.

To reduce the complexity of the detection limit parameter space, I will highlight a particular case study in the following, using a set of simplifiying assumptions. This represents a basic feasibility study and, along with similar case studies, informed the observing time requests of our programme. During a real observing run, achieved detection limits can be quite different, due to: variable observing conditions, weather or technical downtime during parts of the run, changes in target cadence due to telescope pointing restrictions, the presence of red noise, and multi-planet signals. Clearly it is not feasible to simulate such complications *a priori*. For this particular case study, we generated synthetic RV timeseries using the following assumptions.

- Each observation (incl. overhead) lasts 20 min.
- 5 targets are observed at identical cadence.

- A summer night is assumed, with only 400 min (~6.5 hours) of useful observing time. This allows 20 observations per night, i.e. 4 per target at a cadence of 100 min.
- The observing nights are consecutive.
- In a real run, observations are not completely regular in time: a small random jitter was therefore added to the cadence.
- Each observation has a precision of 1 m s^{-1}. Random RV jitter at this level was added to the synthetic timeseries.

This synthetic timeseries was injected with a single Keplerian signal with a particular period and amplitude. Circular orbits were assumed throughout. A simple Lomb-Scargle periodogram search was done on the resulting timeseries. We used the fast period search algorithm described in Press et al. (1992) for this. The injected signal was classified as successfully recovered if the period search returned a matching signal (within 25 % of the injected period) and the analytical False Alarm Probability (FAP) of the detection was below the desired threshold (usually $\leq$ 1 %). This procedure was repeated over a grid of signal amplitudes and periods, covering the parameter space of interest.

The results are presented in Figure 4.9. Each region of parameter space where a signal was successfully recovered is coloured. These detection limits are expressed directly in terms of planetary mass (M$\sin i$) as a function of orbital period, for a set of stellar masses covering spectral types F5 to K5. For reference, the RV semi-amplitude of a 2 $M_\oplus$ planet around a 1 $M_\odot$ star ranges from 2.0 to 0.75 m s^{-1} for circular orbits with periods from 0.25 to 5 days. For a 10 $M_\oplus$ planet around a 1 $M_\odot$ star, the RV semi-amplitudes for the same period-range are 10.2 to 7.5 m s^{-1}. As shown in Fig. 4.9, in this case study we require at least 4 nights to be sensitive to $\lesssim 10 M_\oplus$ planets, i.e. possible rocky bodies. With only 3 nights, no low-mass planet detection could be achieved at 1% FAP,

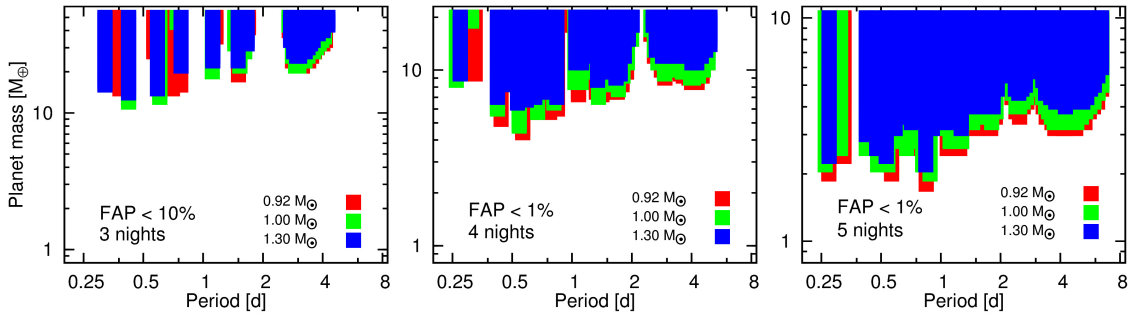

Figure 4.9: FAP detection limits as a function of orbital period for 0.92, 1.0 & 1.3 $M_{\odot}$ stars and observing runs lasting 3, 4 & 5 nights. NB: the planet mass range plotted decreases successively in each panel from left to right. Detection limits for M≤10 $M_{\oplus}$ improve dramatically with 5 nights of observations.

so the 10% case is shown instead. Gaps and irregular edges of the parameter space with successful detections are due to the window-function of the sampling. This can e.g. reduce sensitivity to planetary signals with periods close to (fractional) integer days. Fig. 4.9 demonstrates that a single additional night dramatically pushes down the planet masses to which we are sensitive. This and similar case studies encouragingly showed that the detection of low mass short period planets is feasible in a relatively small number of nights, with an acceptable sample size (5 stars). Using the same assumptions as above, we also simulated detection limits with RV jitter of 10 $m\,s^{-1}$, as might be expected for young, active stars with intrinsic $\log(R'_{HK}) \gtrsim$ -4.7 (c.f. Figure 1.5, Section 1.4.1). In this simulation, 20 - 40 $M_{\oplus}$ planets over the period range 0.25 - 8 d are detected in 5 nights of observations. We typically requested telescope time for 5 consecutive nights, depending on the visibility distribution and number of targets of interest within a given semester / hemisphere.

Fig. 4.10 shows another case study, highlighting the relatively small effect of degrading RV precision from 1 to 2 $m\,s^{-1}$. In this scenario we simulated a run scheduled over 5 consecutive nights that suffers poor observing conditions, with total weather losses of 30%. The number of targets is reduced to

only 2 and 3 in the panels shown, leading to an observing cadence of 40 and 60 min respectively. All other assumptions are as described for the previous case study. Fig. 4.10 indicates that in poor conditions we can achieve detection limits comparable to a "perfect" 5 night run (Fig. 4.9), if the number of targets is reduced. I note at this stage that our actual observations showed the detection

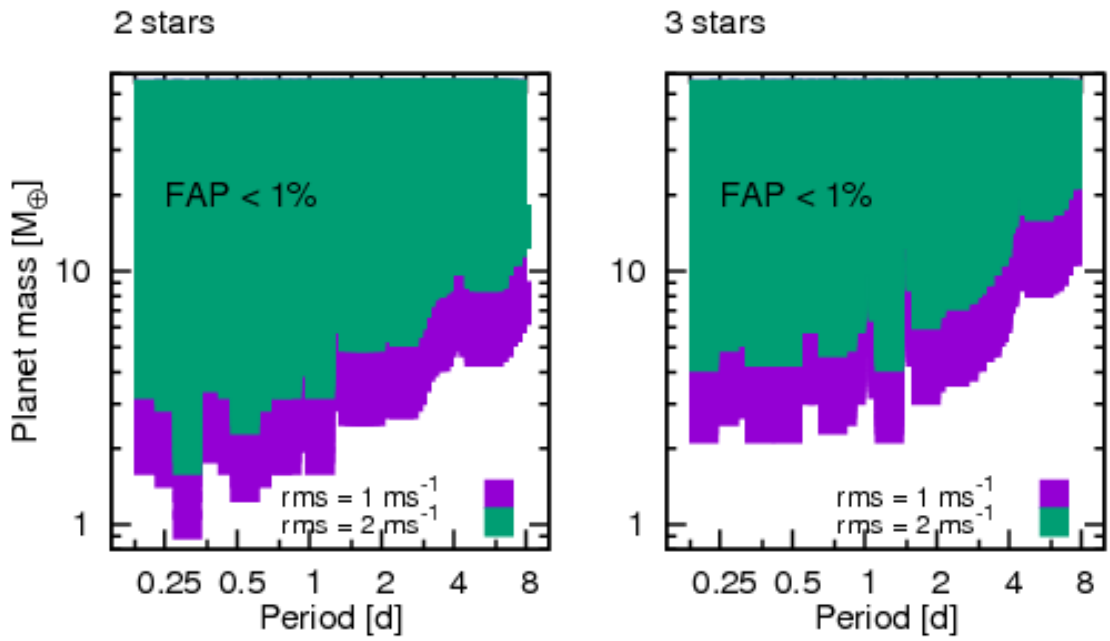

Figure 4.10: 1% FAP detection limits as a function of orbital period for 1.0 $M_{\odot}$ stars and observing runs over 5 consecutive nights, with 3 and 2 targets respectively. To simulate a 30% weather loss, observations were simulated as 1 full night + 1 night gap + 2.5 nights + 0.5 night gap. Using other permutations of the weather loss gaps provided similar results.

limits derived above are too optimistic in terms of the time required for robust planetary detections, due to complications from red noise and especially signal aliasing. Another consideration is that even highly significant signals at periods approaching the baseline of observations typically cannot be claimed as robust detections. More than a single period should be covered to verify the strict periodicity of a Keplerian signal, so a second observing run is typically needed for planet confirmation. These issues will be discussed in detail in Chapter 5.

4.5 A frequentist approach to RV signal detection

In Chapter 5, I will present the search for RV periodicities and statistically significant signals in a frequentist framework.

Historically, the Lomb-Scargle periodogram (Lomb 1976, Scargle 1982) and its variants such as the F-ratio periodogram (Cumming 2004) were widely used to find RV periods in unevenly sampled data. Multi-planet signals were searched in the residual RVs obtained from subtracting successive best fit Keplerians (e.g. Bouchy et al. 2009b, Meschiari et al. 2009, Pepe et al. 2011). To mitigate stellar activity signals, "pre-whitening" was sometimes used (e.g. Queloz et al. 2009, Hatzes 2014). In this approach, Fourier analysis is used to identify sinusoidal (activity-related) RV signals, and they are successively subtracted from the timeseries. Alternatively, linear correlations between RV and activity indicators were fit and subtracted (e.g. Robertson & Mahadevan 2014, Robertson et al. 2014). Planetary signals were sought in the RV residuals obtained after applying these methods. These approaches have been repeatedly shown to be inadequate (Tuomi 2012, Haywood et al. 2014, Anglada-Escude & Tuomi 2015). The fundamental issue is that the parameters describing multiple planet signals and/or RV - activity index trends are correlated (Anglada-Escudé & Tuomi 2012). Combined with residual analysis and/or pre-whitening this can lead to either under- or overestimation of planetary signal significance and false positive or false negative detection claims. False negatives are typically cases where a planet candidate signal is erroneously claimed to be produced by stellar activity (e.g. Robertson et al. 2015 versus Anglada-Escudé et al. 2016b).

In this thesis I use likelihood ratio periodograms, which show the improvement of the likelihood statistic, $\Delta \log L$, of a best fit including a planetary signal compared to the null hypothesis (best fit without a Keplerian signal), at each test period. The natural logarithm of the likelihood L is used for mathematical

convenience. Likelihood periodogram tools for RV analysis were pioneered in Baluev (2009). The likelihood of a hypothesis is the probability of the data for the chosen model. It is the product of the probabilities of each data point observed.

With a suitable definition of the likelihood function (i.e. RV model parameters), this allows global (simultaneous) optimisation of planetary signals, noise parameters (jitter), and linear correlations between RVs and other simultaneous observables. The algorithm we used proceeds recursively, maximising the global likelihood each time an additional planet signal is added to the model. For example, all parameters describing noise, correlations and the first planetary signal are re-optimised when searching for a second planet signal. For context on the mathematical approach for obtaining a best-fit Keplerian signal, see Wright & Howard (2009) and refer back to Section 1.1.1. The recursive approach is first described in Anglada-Escudé & Tuomi (2012), where it was applied in the context of F-ratio periodograms. It was then applied to likelihood periodograms in Anglada-Escudé et al. (2013), and has been used repeatedly for planet discovery (e.g. Anglada-Escude et al. 2014, Tuomi et al. 2014, Anglada-Escudé et al. 2016a). This approach avoids the problems of residual analysis and pre-whitening described above, which can be essential for low RV amplitude, multi-planet systems. The likelihood has a lower intrinsic variance compared with other statistics (such as χ^2) that do not include noise parameterisation (Anglada-Escudé et al. 2016b).

In many cases, planet detection is carried out by combining RV data from a range of instruments. The corresponding subsets have RV zeropoint offsets (c.f. Section 1.1.1) and show different jitter levels above their internal uncertainties. This can be the case even for datasets from the same instrument: as decribed in Section 4.3, major instrumental upgrades can cause zeropoint offsets. In our RV program, targets are often observed at high cadence in multiple

runs, separated by large data gaps of order 1 year. Long-period (planetary or stellar) signals may therefore lead to RV offsets between runs that cannot be well constrained with a Keplerian fit. In such situations, fitting RV offsets between the individual runs can also be advantageous while searching for short-period signals. Note that different jitter values may be appropriate for individual runs even when using the same instrument. For instance, a target star can shift from lower to higher activity phases during the relatively long gap between runs. I will refer to individual sub-datasets as "subsets", regardless of whether they originate from different instruments. I define the ith radial velocity measurement, obtained at time t_i from a data subset as $v_{i,\mathrm{subset}}$. A reference epoch t_0 is (arbitrarily) chosen to be the first observation of the timeseries. The likelihood function L of the observations, following Anglada-Escudé et al. (2013) (see also Anglada-Escudé et al. 2016b), is then given by

$$L = \prod_{\mathrm{subset}} \prod_{i} l_{i,\mathrm{subset}}. \tag{4.4}$$

$$l_{i,\mathrm{subset}} = \frac{1}{\sqrt{2\pi(\sigma^2_{\mathrm{RV},i} + \sigma^2_{\mathrm{subset}})}} \exp\left\{-\frac{1}{2}\,\frac{\epsilon^2_{i,\mathrm{subset}}}{\sigma^2_{\mathrm{RV},i} + \sigma^2_{\mathrm{subset}}}\right\} \tag{4.5}$$

$$\epsilon_{i,\mathrm{subset}} = v_{i,\mathrm{subset}} - (\gamma_{\mathrm{subset}} + \dot{\gamma}\,\Delta t_i + K(\Delta t_i) + A_{i,\mathrm{subset}}) \tag{4.6}$$

$$\Delta t_i = t_i - t_0 \tag{4.7}$$

- $\epsilon_{i,\mathrm{subset}}$ is the ith observed-model residual value. In Equation 4.5, the assumption is that each $\epsilon_{i,\mathrm{subset}}$ value is a Gaussian random variable centred on zero with variance $\sigma^2_{\mathrm{RV},i} + \sigma^2_{\mathrm{subset}}$, where $\sigma^2_{\mathrm{RV},i}$ is the internal uncertainty of the ith measurement (c.f. Section 4.1.2) and $\sigma^2_{\mathrm{subset}}$ represents the excess white noise (jitter) of each subset.

- γ_{subset} is the reference velocity of each data subset.

- $\dot{\gamma}$ is the slope of a linear RV trend, parameterising long-term RV changes that appear linear on the timescale of observations. This can be caused by long-period companions.

- $K(\Delta t_i)$ is the superposition of n Keplerian signals evaluated at Δt_i. Each of these includes five best fit parameters: the orbital period P, the RV semi-amplitude K, the mean anomaly evaluated at reference epoch t_0, M_0, the eccentricity e and the argument of periastron ω (c.f. Section 1.1.1).

- $A_{i,\text{subset}}$ represents a linear correlation with slope C_{subset}, between RV measurements and any other coincident timeseries with values $\xi_{i,\text{subset}}$.

$$A_{i,\text{subset}} = C_{\text{subset}}\, \xi_{i,\text{subset}} \tag{4.8}$$

In typical RV period searches, $\xi_{i,\text{subset}}$ is a stellar activity indicator. In our case, κ values (c.f. Section 4.3.3) can be used alternatively to fit the chromatic systematic effect simultaneously with Keplerian signals. The package we used allows a choice of setting the eccentricity of all Keplerian signals to zero, simplifying the $K(\Delta t_i)$ terms accordingly: the best fit values of P, K, and ω are determined in this case (c.f. Section 1.1.1). It is relatively common practise to force circular orbits in RV analyis when recovering low amplitudes and/or short periods (e.g. Pepe et al. 2011, Gilliland et al. 2013), especially when limited data is available. For the short period orbits of prime interest in our RV program, tidal circularisation is expected to significantly dampen orbital eccentricities over the systems' lifetime (e.g. Jackson et al. 2008). Low eccentricity values are therefore likely, but this can be investigated by comparing maximum likelihood solutions with and without forcing zero eccentricity. The $\dot{\gamma}$ and $A_{i,\text{subset}}$ terms can also be set to zero. This is useful when testing the effect of putative correlations and long-term slopes on Keplerian signal recovery.

The frequentist False Alarm Probability (FAP) of a planetary signal can be computed from its $\Delta \log L$ value. This is done via the likelihood-ratio test, following the algorithms in Baluev (2009) and Baluev (2008) (see Equations 18 and 12 of those studies respectively). The FAP estimates the fraction of times a signal with a given $\Delta \log L$ would be seen in the dataset purely due to unfortunate arrangement of Gaussian noise. There is no universally agreed value in the literature on the FAP theshold for claiming statistically significant planet detections (e.g. Butler et al. 2017). Thresholds of 1% and 0.1% are typically used. Formally significant signals must also be well-constrained in period from above and below - for periods significantly longer than the baseline of observations this will not be the case. Such signals can be formally extremely significant, but nevertheless not constitute detections because a strict periodicity cannot be verified.

In the analysis of Chapter 5, the relative probability associated with Keplerian signals fitted at different periods with high $\Delta \log L$ values in the same likelihood periodogram will be of interest. In such comparisons, the models corresponding to these periods have the same number of underlying parameters. The probability ratio of the observations associated with two models (periods) A and B is then equivalent to the ratio of the likelihoods $\frac{L_A}{L_B}$. This probability ratio can therefore be simply computed via

$$\exp\left\{\Delta \log L_A - \Delta \log L_B\right\} = \exp\left\{\log L_A - \log L_B\right\} = \exp\left\{\log \frac{L_A}{L_B}\right\}. \tag{4.9}$$

4.6 Summary

In this chapter I summarised the most relevant information on the instruments used for the precision RV measurements of this thesis. I included descriptions of the two main analysis techniques (cross-correlation and template-matching), and explained corrections applied to the RV output from automated reduction suites. For the SOPHIE spectrograph these are highly significant in some cases, in particular due to moonlight contamination, and large-amplitude chromatic and CTI effects. I discussed the chromatic effect in detail, and established new empirical CTI corrections for both SOPHIE and HARPS. For the latter instrument, I found that the CTI correction has no significant impact on results. I also gave a general overview of astrophysical (false positive) RV signals that can interfer with planet detection. The simulated detection limits of our high-cadence RV programme are shown to reach 2 $M_{\oplus}$ planets in sub-day periods within five nights of observations, under simplifying assumptions. Finally, I explained the details and advantages of using a recursive likelihood ratio periodogram approach to RV planet detection.

Details on collaborative contributions

My supervisor C.A. Haswell originated the idea of a high precision RV search for short-period low-mass planets orbiting stars with anomalously low activity. Detailed discussions with Z.M. Berdinas and G. Anglada-Escude provided valuable input on measuring and assessing the chromatic effect, including Figure 4.8, and a script outputting κ measurements for all relevant spectra. J.R. Barnes led the analysis of achievable RV detection limits (Section 4.4), generated and analysed the synthetic RV timeseries and provided Figures 4.9 and 4.10. The observing strategies outlined in this Chapter and constraints needed for the simulation of detection limits were a collaborative effort led by me, with input from C.A. Haswell, J.R. Barnes and G. Anglada-Escude. G. Anglada-

Escude provided the algorithms that perform recursive maximisation of the global likelihood and output likelihood ratio periodogram data.

Chapter 5

Radial velocity planet search: Results

In this chapter, I give an overview of the planet search program observations in Section 5.1 and provide an analysis of a systematic effect important for SOPHIE data in Section 5.2. I compare RV results obtained from the DRS and TERRA packages and assess the RV precision obtained with HARPS and SOPHIE. I discuss the RV results for two targets in detail in Sections 5.3 and 5.4.

5.1 Overview of observing runs

Table 5.1 provides an overview of our RV programme up to January 2017. Throughout this thesis, our HARPS runs are abbreviated following the ESO semester nomenclature, with P98 being the most recent semester. In P98, we were awarded two separate runs, P98(A) and P98(B). I will also refer to a small dataset of 12 RVs, "P98(J)", acquired by Co-I James Jenkins during an unrelated observing run as part of a time-sharing agreement. Northern Hemisphere runs using SOPHIE and HARPS-N are referred to using the OPTICON semester nomenclature (2015A etc.). All runs using HARPS and SOPHIE were executed in visitor mode, i.e. with members of our team as the observers. HARPS-N data was acquired in service mode by staff at the TNG.

Table 5.1 shows that the first two observing runs of the programme (P95 & 2015A) suffered very large weather losses. Despite limiting the number of targets in these conditions, we collected insufficient data to allow a conclusive search for periodic RV signals. Only large RV amplitude HJ companions could be ruled out. In both runs, poor seeing and extinction from clouds also degraded the SNR and RV precision reached, in spite of significantly increased exposure times. These targets were therefore re-observed a year later, in the P97 and 2016A runs.

After selection as described in Chapter 2, a total of 22 unique targets have been observed in the programme to date (April 2017). Exactly half of these are located in each hemisphere. For 5 stars, we only obtained very limited, "exploratory" observations (2-14 RVs per target); e.g. during limited visibility windows when primary targets were not accessible. As of January 2017, we have collected more substantial datasets for 17 objects (22-154 RVs per target), which allow period searches for low-amplitude signals.

Table 5.1: Summary of our RV programme. I list the total nights awarded per run. For some runs, these nights were not allocated entirely consecutively and/or additional fractional nights were added via time-sharing agreements with other observers. This is reflected in the run dates. Losses are relative to a run with no downtime and with no significant increases in exposure times in response to poor seeing. The "official" run ID is assigned by the observatories' online archives.

Run	Time awarded [nights]	Instrument	Losses [%][a]	Dates[b]	Targets observed[c]	Official ID	Observing team[d]
P95	5	HARPS	63	2015-09-20→2015-09-24	3	095.C-0799(A)	DS, CAH
P96	3.5	HARPS	0	2015-12-18→2015-12-22	2	096.C-0876(A)	DS
P97	3.5	HARPS	17	2016-07-28→2016-08-03	7	097.C-0390(B)	DS, JD
P98(A)	6	HARPS	0	2016-11-29→2016-12-04	8	098.C-0269(A)	CAH, JC
P98(J)	N/A	HARPS	N/A	2016-12-19→2016-12-26	5	096.C-0499(A)	JJ, MD
P98(B)	6	HARPS	10	2017-01-03→2017-01-08	7	098.C-0269(B)	JRB
2015A	5	SOPHIE	50	2015-04-27→2015-05-01	4	15A.OPT.HASW	DS, CMC, JRB
2015B	5	SOPHIE	15	2015-11-09→2015-11-13	6	15B.DISC.HASW	DS
2016A	7	SOPHIE	35	2016-04-04→2016-04-10	4	16A.OPT.HASW	DS, JC
2016A	1	HARPS-N	0	2016-07-14→2016-07-16	1	OPT16A11	service mode
2016B	5	SOPHIE	9	2016-11-29→2016-12-04	4	16B.DISC.HASW	JRB, JD, GDS

[a] from weather & technical downtime
[b] Start of night UTC
[c] excluding RV standards
[d] D. Staab (DS), J.R. Barnes (JRB), J. Cooper (JC), C.A. Haswell (CAH), J. Doherty (JD), G.D. Smith (GDS), C. McCune (CMC), J. Jenkins (JJ), Matias Diaz (MD).

In most cases, targets have been re-visited across multiple runs to

- increase the significance and robustness of possible detections
- distinguish between alias period solutions
- confirm the strict periodicity of signals with periods similar to the baseline of a single run
- clarify the effect of possible stellar, non-Keplerian signals and
- constrain high-amplitude signals from long period (> 1 year) companions.

These issues are discussed in more detail in Sections 5.3 and 5.4. It is beyond the scope of this thesis to provide a full analysis of each of the targets' data. Active follow-up observations of many targets are still ongoing at the time of writing, for the reasons listed above. Preliminary results for several targets remain "confidential" until they are ready for publication as either conclusive planet detections or null results with detection limits. I will discuss general results on systematic effects and RV stability seen in our datasets, followed by key results for 2 targets. Note that publications are in preparation for these objects.

5.2 Analysis of the chromatic effect

5.2.1 SOPHIE data

As already illustrated by Figure 4.8 in Section 4.3.3, our SOPHIE data clearly showed variable pSED slopes due to chromatic flux losses. In this Section, I will show the impact this has on the RV and FWHM values measured with SOPHIE spectra. This work was inspired by the B16 publication, and constitutes the first analysis of the chromatic effect applied to SOPHIE that I am aware of. To recap Section 4.3.3, if SOPHIE were to mirror the HARPS-N results in B16, we would expect FWHM - κ and RV - κ correlations. Because the SOPHIE DRS does not apply a flux-weighting correction before CCF computation, while TERRA includes flux-normalisation, we primarily expect the CCF-derived RVs to be affected. Using weighted least squares (WLS) fitting via the STATSMODELS package (Seabold & Perktold 2010), I tested the significance of correlations between RV, FWHM and κ and measured the slopes of the corresponding linear fits (Table 5.2, and Figures 5.1, 5.3, 5.6). I used the p-value associated with the F-test included in the WLS routine from STATSMODELS. The analysis in this Section is focussed on the datasets of six targets from our 2015B run with more than 10 observations.

FWHM

For all targets, I observed very strong correlations between κ and FWHM (Fig. 5.1). **This indicates that the chromatic effect observed by B16 for HARPS-N affects SOPHIE data.** Figure 5.2 illustrates the nightly trends of the chromatic effect over the course of an observing run by plotting the FWHM, κ, and airmass of one of our targets. On the last night of this run, poor seeing acted as a "natural scrambler": the overall flux losses at the fibre entrance were larger, but less chromatic. Therefore the amplitude of the systematic effect is diminished.

Table 5.2: Results of WLS fitting of the chromatic effect on FWHM and RV from SOPHIE 2015B data, and the RV variability observed for each target without de-trending. Note that HD 30708 was observed as an RV standard.

Target	SpType	FWHM-κ slope [m s$^{-1}\mu$ counts^{-1}]	p-value	DRS CCF RV-κ slope [m s$^{-1}\mu$ counts^{-1}]	p-value	TERRA RV-κ slope [m s$^{-1}\mu$ counts^{-1}]	p-value	DRS RV RMS [m s^{-1}]	TERRA RV RMS [m s^{-1}]
HD 184960	F5	36.1 ± 5.2	1.6x10^{-7}	-0.8 ± 1.8	0.65	-0.2 ± 1.9	0.93	3.3	3.3
HD 35768	F8	40.0 ± 2.8	8.1x10^{-10}	-1.2 ± 0.7	0.10	-1.2 ± 0.6	0.07	3.7	3.2
HD 30589	F9	30.4 ± 1.9	2.2x10^{-13}	-0.9 ± 0.9	0.32	0.3 ± 0.7	0.70	6.5	5.1
HD 19902	G5	35.6 ± 1.7	3.4x10^{-20}	-1.8 ± 0.7	0.03	0.6 ± 0.9	0.50	6.2	6.2
HD 30708	G5	21.4 ±0.8	4.1x10^{-20}	-1.6 ± 0.4	0.001	-0.2 ± 0.3	0.59	3.0	1.6
HD 11130	G9	28.4 ±1.0	6.4x10^{-24}	-0.4 ± 0.3	0.12	-0.6 ± 0.3	0.04	2.3	2.2

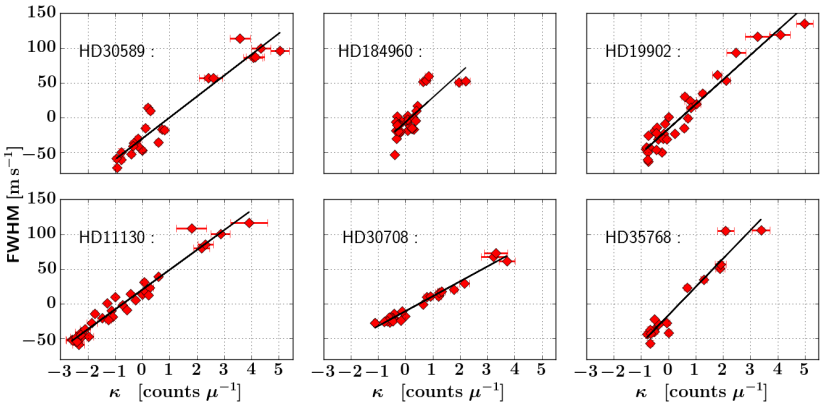

Figure 5.1: Observed κ - FWHM correlations with best fit results from WLS (see Table 5.2) for SOPHIE 2015B data. The same axis ranges are used throughout to visualise the correlations consistently.

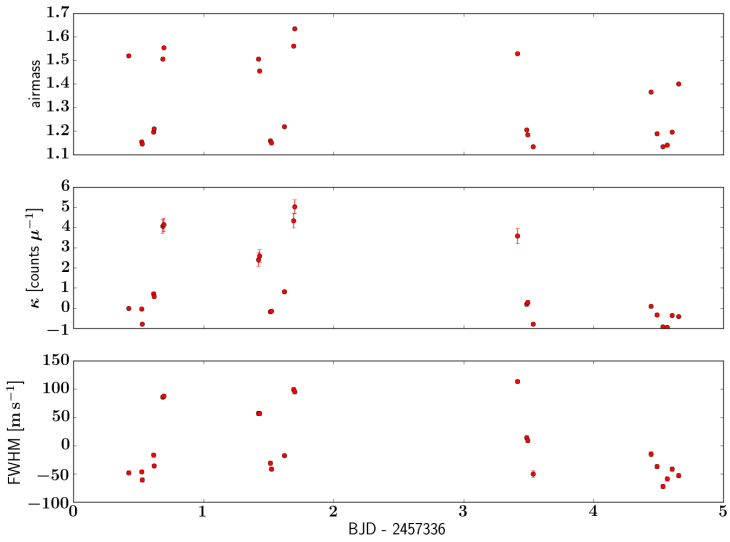

Figure 5.2: SOPHIE timeseries of HD30589. The chromatic effect, parameterised by κ, follows nightly airmass trends, as does the FWHM.

DRS RVs

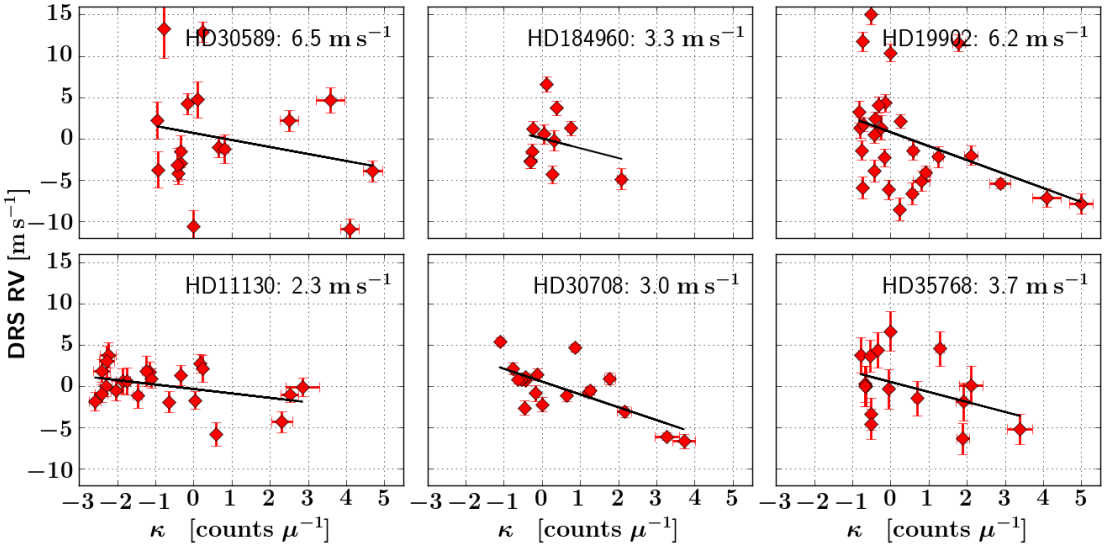

Figure 5.3: Observed κ - DRS RV correlations and linear fits from WLS (see Table 5.2). Note consecutive short exposures were binned into ~15 blocks to remove asteroseismic p-mode noise and to compare all targets more consistently (c.f. Section 4.3.2). To better visualise the divergent RV variability of our targets, the same axis ranges are used throughout. The RV RMS (without de-trending) is shown for each target.

Fig. 5.3 and Table 5.2 strongly suggest that the DRS RVs of several targets are influenced by the chromatic effect, with additional astrophysical variability at varying levels. Given the results for the FWHM, it seems likely that all RVs are systematically affected, but that for some targets astrophysical RV variability dominates. Figure 5.4 shows the κ and DRS RV timeseries of 2 targets from our 2015A run. Since these stars have similar coordinates they were always observed at comparable airmass, leading to matching κ values. This is a useful illustration of how the chromatic effect can cause correlated intra-night DRS RV variability at the ~15 m s^{-1} peak-to-peak level. The RMS of the two datasets is 6.5 and 8.2 m s^{-1} respectively, well above SOPHIE's typical instrumental stability. Note that our RV analysis (including data from other runs) revealed no significant stellar or planetary signals for these targets. It is important to emphasise that this chromatic effect is a particular concern for our

observing strategy aimed at finding ultra-short period planets. We observe targets at high cadence, over relatively large airmass ranges. Typical RV surveys observe a target only once or twice per night near zenith. Therefore a much small κ range is likely sampled, greatly reducing the importance of the effect.

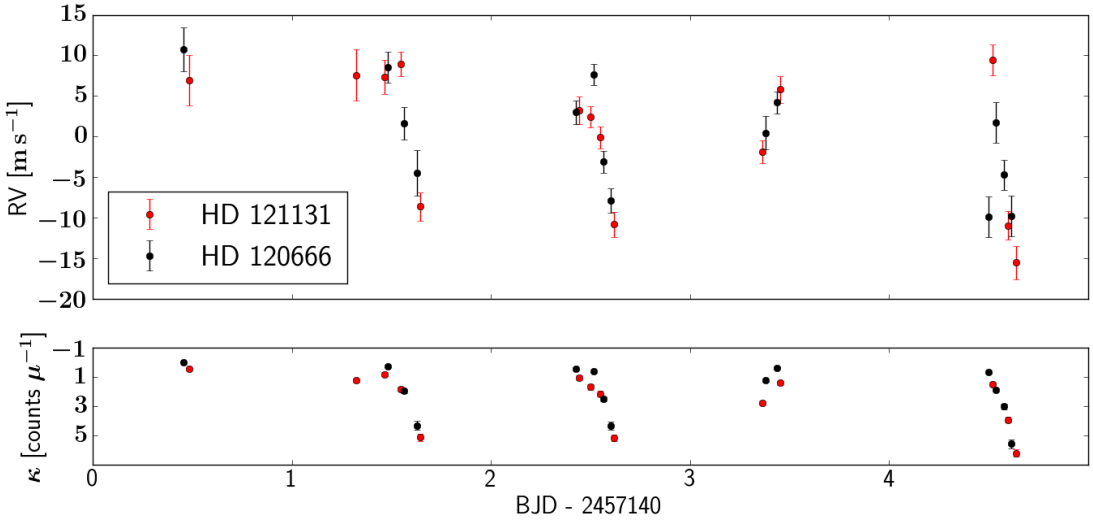

Figure 5.4: DRS RV timeseries of HD 120666 & HD 121131 (derived with a G2 CCF mask), and the corresponding κ values. For illustration purposes, the y-axis of the lower panel is inverted. The co-trending variability of the targets is due to the chromatic effect.

The fitted slope values in Table 5.2 vary significantly, in particular for the FWHM-κ correlation. This indicates a dependence of the chromatic effect on additional parameters such as the spectral type. I also tested the effect of changing the CCF mask for the same dataset. For the K1 star HD 121131 (observed in 2015A) using the G2 and K5 masks gives significantly different slopes (Table 5.3). A thorough investigation of the dependence of the chromatic effect on stellar parameters and choice of CCF mask is left for future work. This may become redundant if TERRA RVs are used and/or the SOPHIE DRS is updated (see below).

Table 5.3: The chromatic effect is clearly present in the 2015A CCF data of the K1 star HD121131. Its RV amplitude (slope) increases dramatically when the non-optimal G2 CCF mask is used in the DRS computation.

Target	FWHM-κ slope [$ms^{-1}\mu$ $counts^{-1}$]	p-value	RV-κ slope [$ms^{-1}\mu$ $counts^{-1}$]	p-value
HD 121131 (K5 mask)	33 ± 2	$2x10^{-9}$	-1.1 ± 0.5	0.04
HD 121131 (G2 mask)	17 ± 1	$7x10^{-9}$	-4.3 ± 0.5	$8x10^{-7}$

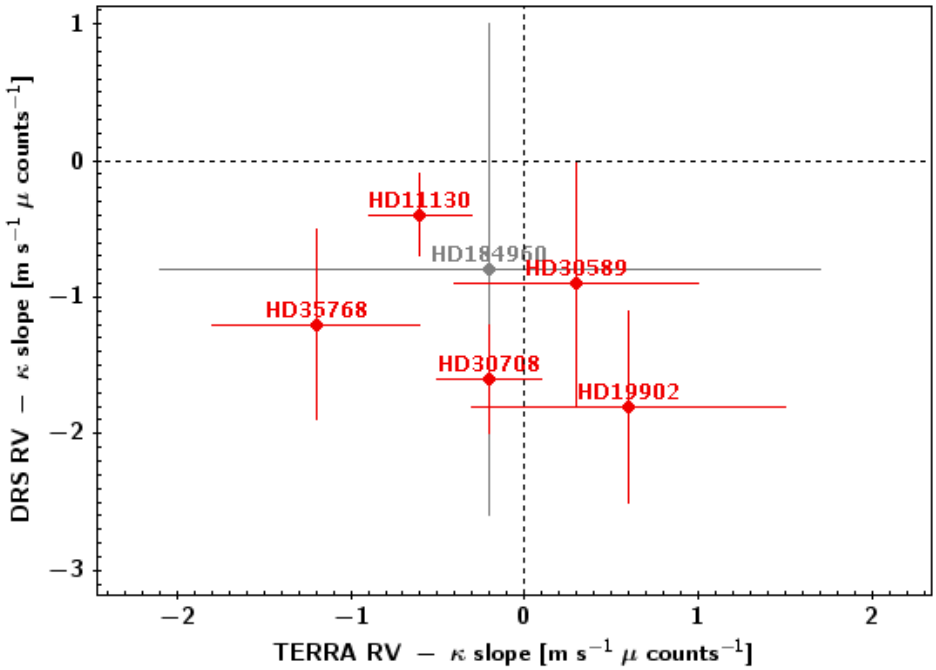

Figure 5.5: Illustration of the WLS results in Table 5.2 for the SOPHIE chromatic effect. For each 2015B target, the best fit TERRA RV-κ slope is plotted against the corresponding DRS RV-κ value. TERRA slope values are clustered around zero, while DRS values are at the -1 m $s^{-1}\mu$ $counts^{-1}$ level.

TERRA RVs

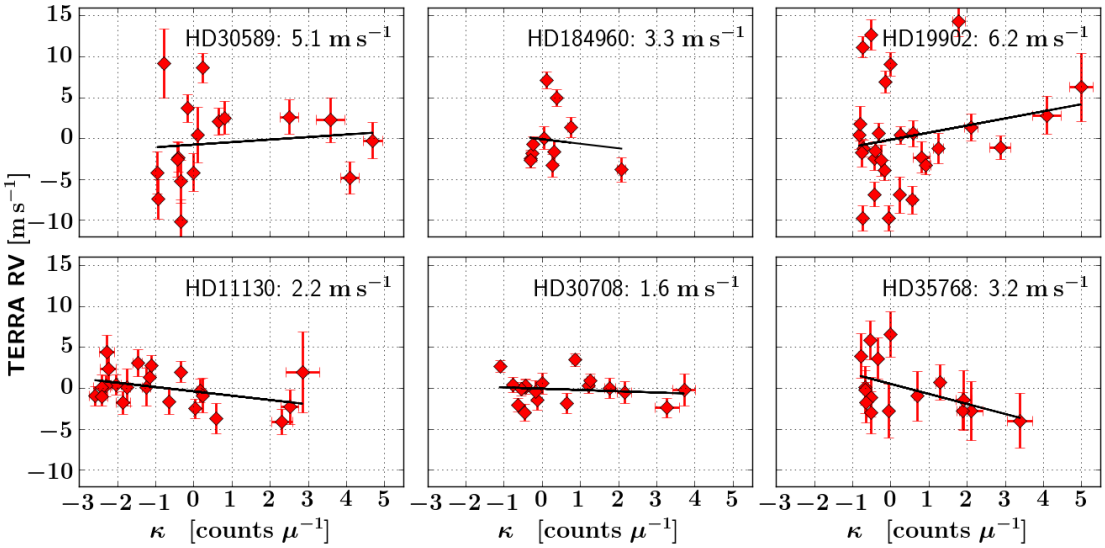

Figure 5.6: Same as Figure 5.3, but for RVs reduced with the TERRA package. Note that HD 30708 was observed as an RV standard, and a clear null result with respect to Keplerian signals was obtained for HD 11130 (see below).

Two targets (HD11130 and HD35768) show evidence for a non-zero chromatic slope in their TERRA RVs (Figure 5.5 and Table 5.2). However the TERRA timeseries with by far the lowest RV scatter (RV standard HD30708) provides no evidence for a TERRA chromatic effect, unlike its DRS counterpart. Figure 5.5 highlights the contrasting chromatic behaviour of TERRA and DRS RVs for all targets. Overall there is no compelling case for a consistent systematic effect for TERRA data, while DRS products are clearly affected. While a low-level TERRA chromatic effect certainly cannot be ruled out with current data, the observed RMS of the RV standard HD30708 and the null result target HD11130 are consistent with the expected SOPHIE stability of 1-2 m s^{-1}.

Consequences for period detection

I highlight an example of the impact of the chromatic effect on likelihood periodograms of DRS RV and FWHM in Figure 5.7. All prominent RV peaks

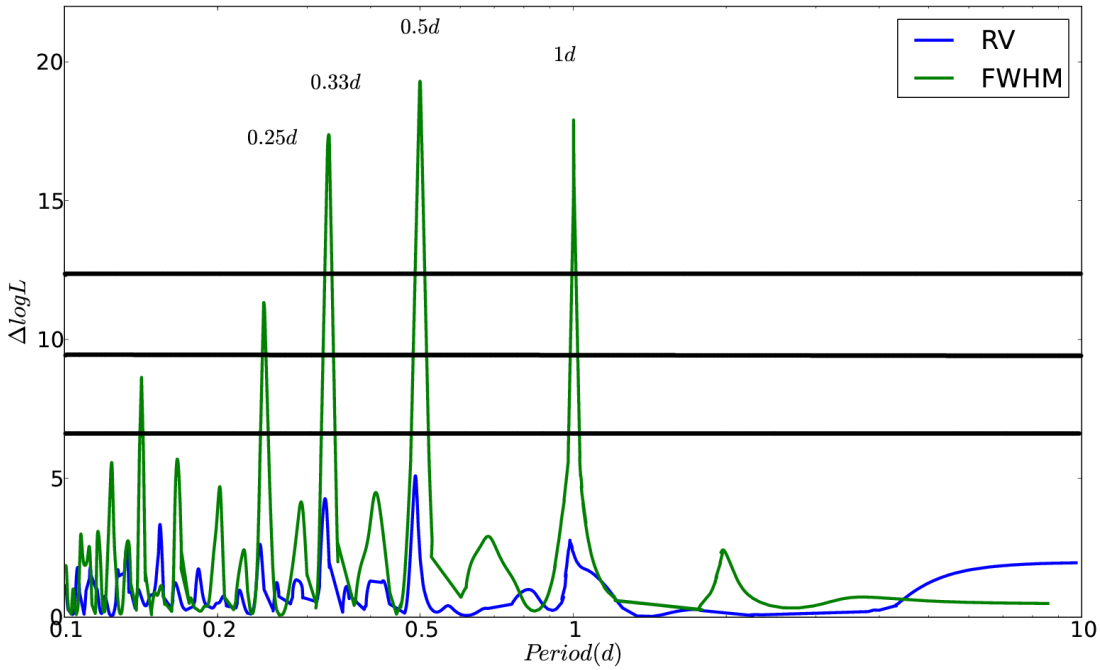

Figure 5.7: Likelihood periodograms for the DRS RV and FWHM timeseries of HD 11130, highlighting the chromatic effect. None of the RV peaks is formally significant, in contrast to the prominent FWHM peaks at (fractional) integer days. The 10%, 1% and 0.1% FAP thresholds are shown.

are located where the FWHM also shows periodicities. The FWHM peaks are found at frequencies of $1d^{-1}$, $2d^{-1}$, $3d^{-1}$, etc. and $1.5d^{-1}$, $2.5d^{-1}$, $3.5d^{-1}$, etc. This is the chromatic systematic effect, with periodicity appearing at fractional days due to nightly airmass trends. The RV signal at 0.5 d has a best fit semi-amplitude of 3.1 m s^{-1} and is highly likely to arise from the chromatic systematic effect, given the behaviour of the FWHM. I find no formally significant RV peaks: their analytical FAP is > 10%. Given that the most prominent RV peak positions are consistent with a systematic effect, there is no evidence whatsoever for Keplerian signals. When fitting a simultaneous correlation with κ, the 0.5d signal decreases in significance to 80% FAP. No signals away

from the fractional day aliases increase in significance, further evidence against planetary signals. This target provides a null result for our planet search and evidence for the chromatic effect on DRS RVs. Although the RV-κ correlation has only marginal significance in this case (Table 5.2), the periodogram clearly indicates a low-level impact on the RVs.

Summary: SOPHIE data

The chromatic effect impacts RVs from CCF computation with the current SOPHIE DRS at the $\sim$1 m s^{-1} μ counts^{-1} level (Fig. 5.5). A key result is that the nominal SOPHIE precision may be dramatically degraded from 1-2 m s^{-1} to $\sim$7 m s^{-1}, when observing targets over moderate airmass ranges and if this systematic effect is not accounted for. In the DRS RV timeseries of high-cadence runs, clear nightly trends up to $\sim$15 m s^{-1} peak-to-peak are seen. In the same circumstances FWHM values are dramatically affected, with $\sim$150 m s^{-1} peak-to-peak amplitudes.

I showed that there is no clear evidence for a chromatic effect in TERRA RV data obtained from SOPHIE. This indicates that the dominant component of the SOPHIE chromatic effect can be addressed by flux normalisation (c.f. Section 4.3.3) before CCF computation, as is the case for HARPS-N (B16). The SOPHIE DRS should be upgraded to implement this step.

A common linear fit to detrend all RV data is clearly not justified *a priori*, given the results in Table 5.2, and since such pre-whitening of RV data is generally unadvisable (Section 4.5). The more robust approach is global optimisation, i.e. simultaneously fitting Keplerian signals and a RV-κ linear correlation term. The significance of real Keplerian signals in the data should increase when such a correlation term is included in the fit, assuming the RV-κ trend is affecting the relevant dataset. Comparing results with and without a correlation term is

therefore a useful diagnostic.

The chromatic effect greatly reduces the usefulness of the SOPHIE FHWM as a tracer of real stellar phenomena, whenever targets are observed over a wide airmass range. To make use of the FWHM as a diagnostic of real stellar signals at the 100 $\mathrm{m\,s^{-1}}$ level for data with a large κ range, the SOPHIE DRS must be upgraded to use flux normalisation. On the other hand, any spurious periodic signal seen in the RVs that is driven by the chromatic effect should be recovered more clearly in the FWHM periodograms. Since planetary signals do not affect the FWHM, this parameter can be used as a valuable diagnostic to distinguish spurious chromatic signals from planet candidates.

For TERRA RVs, SOPHIE instrumental stability at the expected 1-2 $\mathrm{m\,s^{-1}}$ level was achieved in our high-cadence observations over a few nights. There is a clear case for using only TERRA RVs in the analysis of our SOPHIE data, as the relatively large amplitude chromatic effect can be avoided. For the datasets in Table 5.2, the RMS of TERRA RVs is reduced by up to 1.4 $\mathrm{m\,s^{-1}}$ compared to DRS values.

5.2.2 HARPS data

Chromatic losses and variable pSED slopes were also seen in our HARPS data. The key difference to SOPHIE is that the HARPS DRS implements flux-weighting for all spectral types except M-dwarfs (Section 4.3.3). Given the results from B16 and Section 5.2.1, we therefore expect the most significant (if any) chromatic effect in M-dwarf CCF data. While our science targets are FGK stars, we observed the low-activity M1 star GJ191 ("Kapteyn's star") as an RV standard during our programme. This star's velocities display somewhat controversial, long period ($P = 49$ d; 121 d), low-amplitude ($K = 2.3\,\mathrm{m\,s^{-1}}$) planetary signals (Anglada-Escude et al. 2014 versus Robertson et al. 2015 versus Anglada-Escudé et al. 2016b). On the ~5 night timescales of our runs GJ191 is very stable. The best fit solution of Anglada-Escude et al. (2014) shows a HARPS RV jitter of only $0.65\,\mathrm{m\,s^{-1}}$.

In the following I repeat analysis of the chromatic effect as done for SOPHIE and provide WLS results for correlations between RVs, FWHMs and κ via the STATSMODELS package, including the F-test p-value. I used the datasets of our P96 and P97 targets with more than 14 observations. This criterion also excludes stars with high amplitude astrophysical variability, e.g. HD11231 (c.f. Section 5.3). Figure 5.8 illustrates a clear chromatic effect for the GJ191 FWHM values. The only other FWHM-κ correlation with a p-value below 10% occurs for HD2134 (Table 5.4). This is driven entirely by the outlier point seen in Figure 5.8: without this value there is no significant correlation (p-value of 0.32). Clearly, there is no consistent evidence for a chromatic effect in the FWHM our FG-type stars.

Table 5.4: Results of WLS fitting of the chromatic effect on FWHM and RV from HARPS P96-97 data, and the RV variability observed for each target without de-trending. Note that GJ 191 was observed as an RV standard.

Target	SpType	FWHM-κ slope [m s$^{-1}\mu$ counts^{-1}]	p-value	DRS CCF RV-κ slope [m s$^{-1}\mu$ counts^{-1}]	p-value	TERRA RV-κ slope [m s$^{-1}\mu$ counts^{-1}]	p-value	DRS RV RMS [m s^{-1}]	TERRA RV RMS [m s^{-1}]
GJ 191	M1	-20.4 ± 0.5	1.1x10^{-24}	-0.3 ± 0.4	0.40	-0.2 ± 0.2	0.43	1.3	0.8
HD 28471	G3	-0.04 ± 0.4	0.92	-0.3 ± 0.4	0.43	-0.4 ± 0.4	0.33	1.2	1.0
HD 38677	F9	-1.5 ± 1.3	0.24	-0.9 ± 1.3	0.48	-1.4 ± 1.3	0.32	2.3	2.3
HD 2134	G7	-3.0 ± 1.4	0.05	-0.7 ± 0.8	0.35	-1.3 ± 0.7	0.09	1.3	1.3
HD 200133	G2	-0.5 ± 2.0	0.78	1.0 ± 1.0	0.37	1.3 ± 0.8	0.14	2.1	1.8

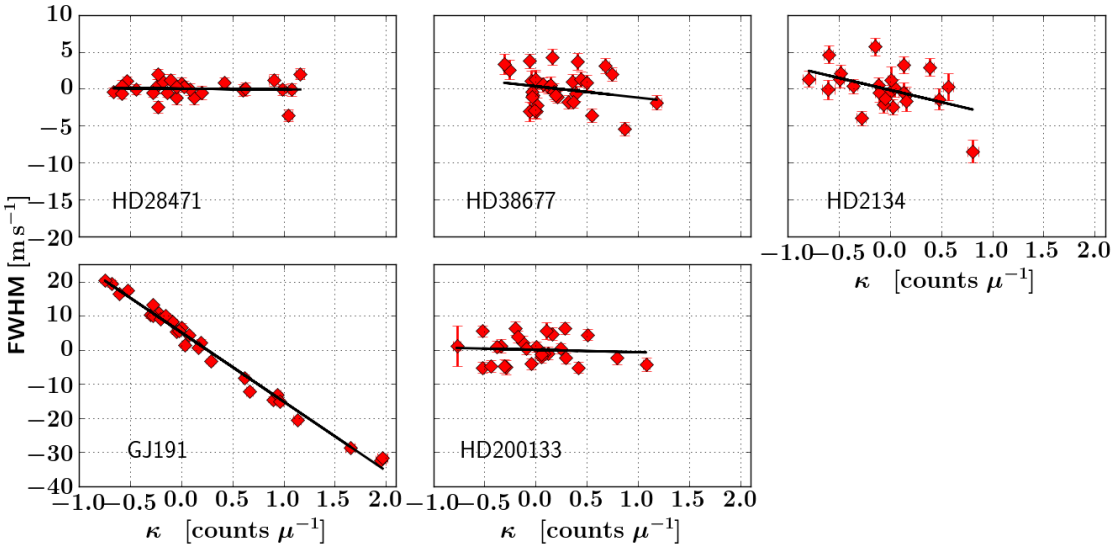

Figure 5.8: Observed κ - FWHM correlations with best fits from WLS (see Table 5.4) for HARPS P96-97 data. Note the change in axis ranges between top and bottom rows, and compared to the corresponding SOPHIE plot (Fig. 5.1)

As seen in Figures 5.9 and 5.10 and Table 5.4 there is also no evidence for the chromatic effect in both DRS and TERRA RVs for any of our targets. The results in Table 5.4 for our stars with the lowest TERRA RV variability, HD28471 and GJ191, indicate that any putative chromatic effect on the RVs is not relevant at the instrumental precision level in our datasets. As expected, TERRA RVs show less scatter than DRS values for most of our targets, with RMS reductions between 0.2 and 0.5 m s^{-1} (c.f. Section 4.1.3). The lowest TERRA RV jitter values reached are 0.8 and 0.5 m s^{-1} for HD28471 and GJ191 respectively. This demonstrates that the HARPS instrumental stability was well below 1 m s^{-1} during the P96 run when this data was collected. Note that for the same datasets, the DRS CCF jitter is 1.1 and 1.0 m s^{-1}.

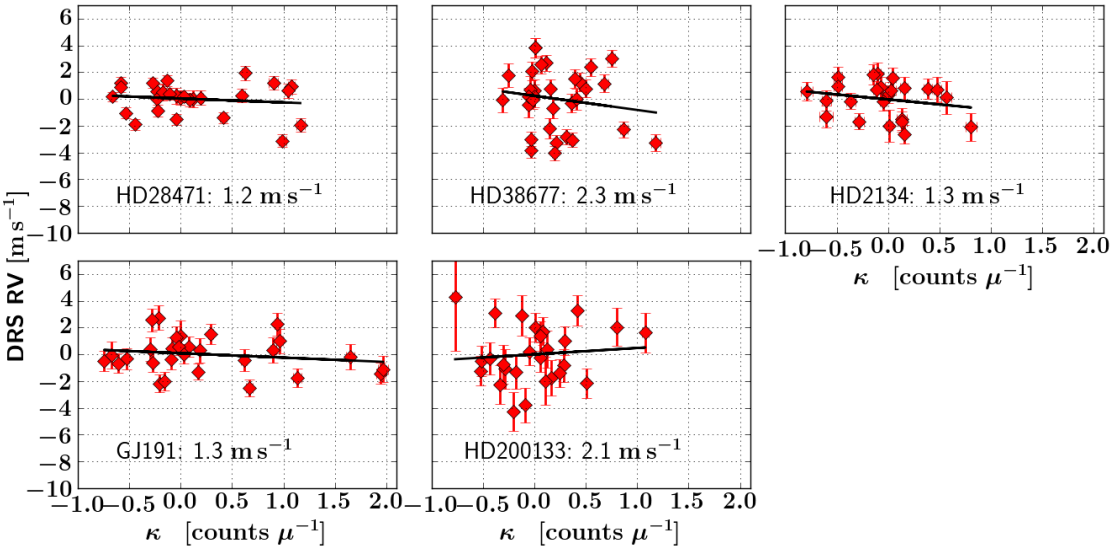

Figure 5.9: Observed κ - DRS RV correlations with linear fits from WLS for HARPS P96-97 data. The RV RMS (without de-trending) is shown for each target.

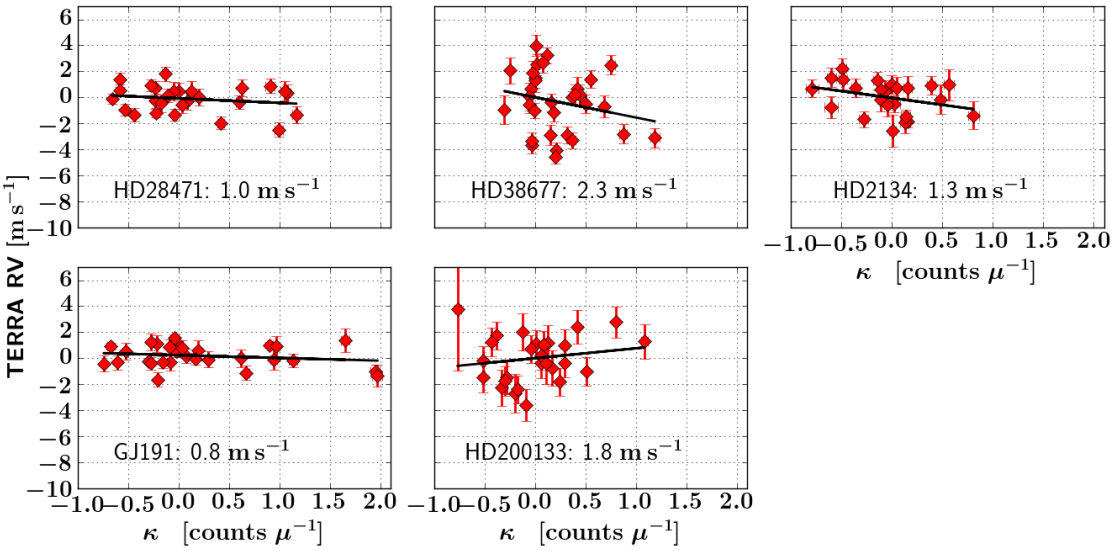

Figure 5.10: Same as Figure 5.9 for TERRA RVs.

Summary: HARPS data

In our HARPS data, the only conclusive impact of the chromatic effect is on the FWHM of an M-dwarf RV standard. The RVs of this target showed exquisite stability and no correlation with κ. I found no convincing evidence that the FWHM of our FG-type targets and the HARPS RVs of any of our targets are impacted by the chromatic effect. The RV jitter levels measured confirm that excellent HARPS instrumental stability (< 1 m s^{-1}) can be achieved on the timescale of our runs while observing targets at high cadence over a relatively large airmass range. **Since TERRA provides more precise RVs than the DRS for several targets, I will use TERRA RVs throughout the remainder of this thesis.**

5.3 A Saturn-mass companion orbiting the pulsating star HD11231

5.3.1 Stellar properties and archival data

HD11231 is a bright (V = 8.6) F5V star that has not been studied in detail in the literature to date. This target was flagged up as an activity outlier in our MS sample due to its single P13 value of $\log(R'_{\rm HK})$ = -5.26; compiled from Arriagada (2011). This is the median activity level over seven measurements with a baseline of 813 days. All datapoints were consistently below the basal level with -5.17 > $\log(R'_{\rm HK})$ > -5.36 (Arriagada, private communication). We therefore prioritised this target for our first HARPS observing run.

No spectroscopic analysis of HD11231 has been published to date. We used the high SNR co-added spectrum from our 49 HARPS observations (Section 5.3.2) to derive the stellar atmospheric parameters given in Table 5.5. Details of the methodology for deriving these parameters are described in Ryabchikova et al. (2009) and Fossati et al. (2010b, 2011). In brief,

- The effective temperature ($T_{\rm eff}$) was obtained by simultaneous analysis of hydrogen (Hα and Hβ) and metal lines (Ti I, Ti II, Cr I, Fe I, Fe II, and Ni I). Synthetic spectra were fit to the hydrogen line profiles, and the excitation equilibrium was imposed for the metallic lines.

- The surface gravity was derived from fitting the gravity-sensitive Mg I b lines, and imposing the ionisation equilibrium for Ti, Cr, and Fe.

- The projected rotational velocity ($v \sin i$) and the macroturbulent velocity ($v_{\rm mac}$) were measured by fitting synthetic spectra to 120 weakly blended lines. The relatively high value of $v_{\rm mac}$ = 8.6±1.0 km s^{-1} is not unusal, given the other stellar parameters (c.f. Grassitelli et al. 2015).

SED analysis of HD11231 was carried out by fitting synthetic fluxes, calculated with the derived atmospheric parameters, to the observed Johnson (Kharchenko et al. 2009), 2MASS (Cutri et al. 2003), and WISE (Cutri & Al. 2012) photometry. The photometry was converted to physical units using the calibrations in Bessell et al. (1998), Van Der Bliek et al. (1996), and Wright et al. (2010). This analysis was combined with the measured parallax (Table 5.5), to estimate the stellar radius (R_*) and interstellar reddening, $E(B-V)$. Combining R_* and $T_{\rm eff}$ values gives a stellar luminosity L_* of 1.41±0.16 $L_\odot$.

Finally, the PARAM tool (da Silva et al. 2006) was used for isochrone fitting. The input parameters were the effective temperature, metallicity, V-band magnitude, and parallax, along with a Kroupa mass function (Kroupa 1998), assuming a constant star formation rate. This analysis showed a stellar age of 2.0±0.3 Gyr, a mass of 1.44±0.03 $M_\odot$, a $\log g$ value of 4.10±0.04, and a radius of 1.78±0.09 $R_\odot$. This is a good match to the results from the spectral and SED analyses (Table 5.5).

Archival RV data

We were given access (Arriagada, private communication) to the seven archival RV measurements corresponding to the activity data reported in Arriagada (2011). These RVs are from the Magellan Planet Search Program (Minniti et al. 2009), using the MIKE spectrograph at the Las Campanas Observatory (Bernstein et al. 2003) and have not been published elsewhere to date. Note that MIKE showed instrumental stability at the 5 m s^{-1} level (López-Morales et al. 2008). As illustrated by Figure 5.11 this sparsely sampled timeseries shows significant RV variability; in one instance at the 50 m s^{-1} level over 1 day. The RMS of the MIKE dataset is 23 m s^{-1}, more than 5 times larger than the internal uncertainties of 4 m s^{-1} for this star. If the instrumental RV jitter is added in quadrature, the uncertainties are 6 m s^{-1}.

Table 5.5: HD11231 stellar parameters. Atmospheric parameters from our spectral and SED analysis are included in the top section, while our isochrone fitting results are listed below.

Parameter	value	Reference
Spectral Type	F5V	Hipparcos[a]
Parallax [mas]	7.45 ± 0.24	TGAS[b]
Distance [pc]	134 ±4	TGAS
V [mag]	8.6	Hipparcos
$B-V$ [mag]	0.463 ± 0.015	Hipparcos
$E(B-V)$ [mag]	0.026 ± 0.008	this work
$\log(R'_{\rm HK})$	-5.26	Arriagada 2011
$T_{\rm eff}$ [K]	6500 ± 100	this work
$\log g$ [cm s^{-2}]	3.8 ± 0.2	this work
$v \sin i$ [km s^{-1}]	5.1 ± 1.0	this work
$v_{\rm mac}$ [km s^{-1}]	8.6 ± 1.0	this work
$v_{\rm mic}$ [km s^{-1}]	1.6 ± 0.1	this work
R_* [R$_\odot$]	1.87 ± 0.09	this work
L_* [L$_\odot$]	1.41 ± 0.16	this work
M_* [M$_\odot$]	1.44 ± 0.03	this work
$\log g$	4.1 ± 0.04	this work
R_* [R$_\odot$]	1.78 ± 0.09	this work
L_* [L$_\odot$]	1.27 ± 0.16	this work
Age [Gyr]	2.0 ± 0.3	this work

[a]XHIP compilation; Anderson & Francis 2012
[b]Gaia Collaboration et al. 2016

Clearly a dataset of only 7 observations does not allow any real analysis, but the high amplitude RV variability seen in Fig. 5.11 further motivated our high-cadence follow-up.

5.3.2 Observations

We acquired a total of 49 RV measurements over four observing runs with HARPS (Figure 5.12): 17 points in P95; 13 points in P97; 13 points in P98(A) & 6 points in P98(B). Initial high cadence observations in P95 showed the RVs were dominated by variability of at least ~20 m s^{-1} (peak-to-peak) on timescales > 1 day. High weather losses in this run (Table 5.1) prevented robust detection of well-sampled periods on such timescales. In later runs we observed at lower

cadence, with 1-3 RVs per night, matching the timescale of RV variability. Due to a time-share agreement in P97 we were able to follow this target over a 7 night baseline. We continued observations throughout P98, to confirm the RV

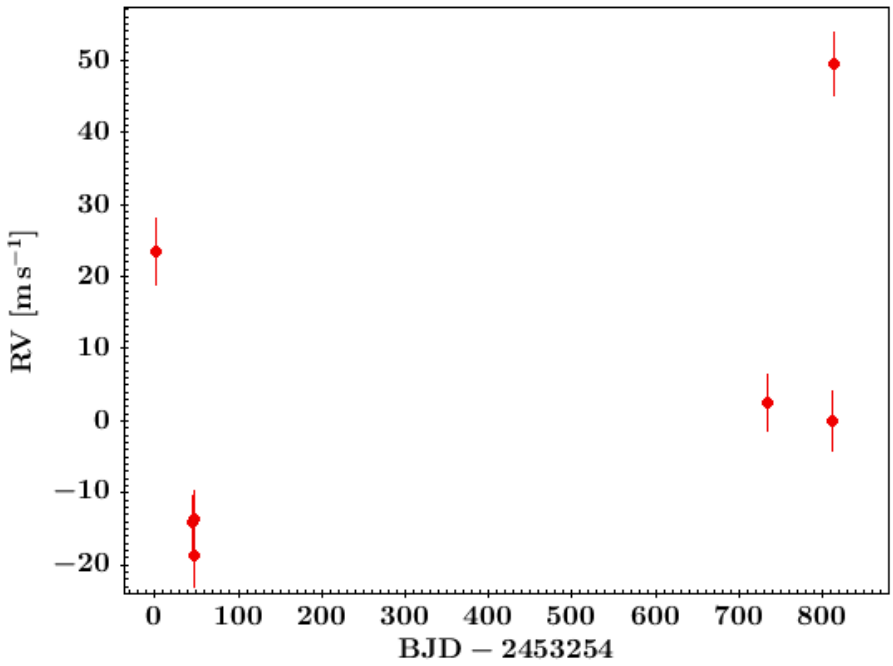

Figure 5.11: Archival MIKE observations of HD11231

periodicity detected in the P95+P97 dataset (Section 5.3.3) and to monitor the prominent stellar line profile changes seen for this target. HD11231 showed the highest amplitude short-period RV variability of all our HARPS targets: up to 150 m s^{-1} shifts occurred over 1 day, and the dataset's RMS was 30 m s^{-1}.

Initially exposure times were varied between 900 and 1800 seconds, depending on weather conditions. In later observing runs, exposure times were reduced to 10 minutes in 3 instances. The corresponding small increase in stellar jitter contribution from p-mode oscillations (Sec. 4.2) is negligible compared to the 150 m s^{-1} peak-to-peak variability of this target. Variations in extinction and seeing across our observing runs resulted in a relatively wide range of SNR (50-160) and σ_{RV} values (1-4 m s^{-1}, with a median of 2.0 m s^{-1}). Note the RV precision was reduced by line broadening compared to stars with lower macroturbulent broadening (Table 5.5).

5.3.3 RV periodicity

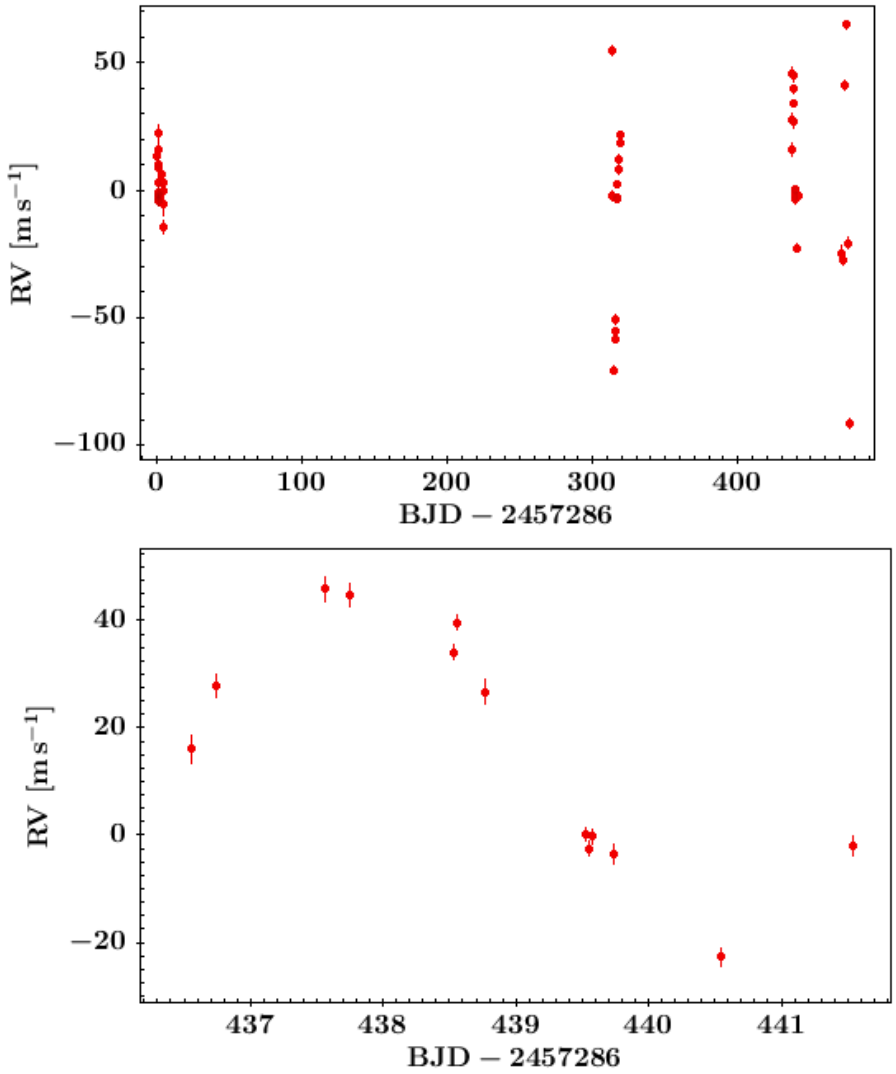

Figure 5.12: HARPS timeseries of HD11231 for all observing runs (top) and P98(A) only (bottom). In several cases uncertainties are smaller than the symbol sizes.

As seen in Figure 5.12, our HARPS timeseries is dominated by variability within each run, with no clear RV offsets between them. I therefore searched for RV periods treating all HARPS data as a single *subset* in the likelihood model (c.f. Section 4.5). The top panel of Figure 5.13 shows the likelihood periodogram obtained with this RV model. The very high significance signal detected initially in P95+P97 around P=6 d was confirmed by adding P98 data: its analytic FAP was boosted to 6.4 $\times 10^{-5}$ %.

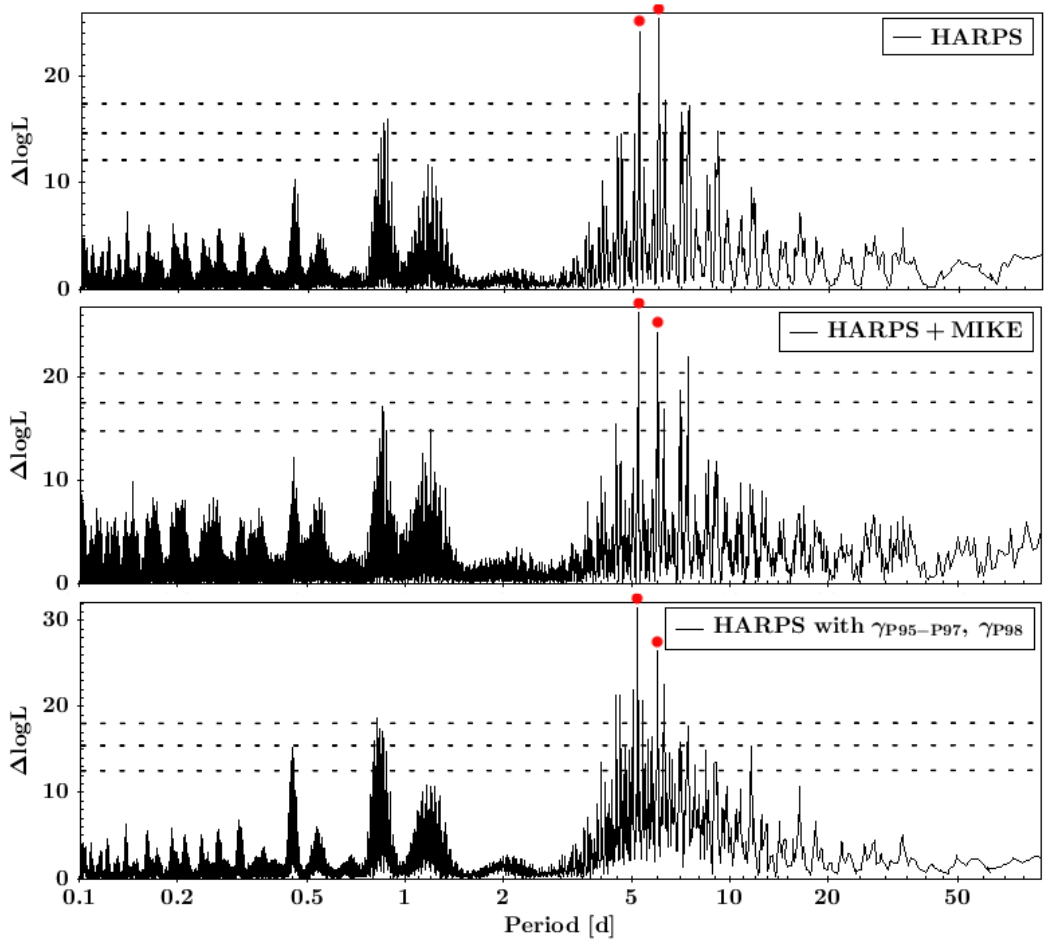

Figure 5.13: HD11231 RV likelihood periodograms for our HARPS data (top), the combined HARPS and archival MIKE data (middle) and the HARPS data when adding a zeropoint offset between the P95-P97 and the P98 subsets (bottom). The 5.998 d and 5.206 d peaks are highlighted, and the 10%, 1% and 0.1% FAP thresholds are shown.

The prominent periodogram peaks seen either side of 1 d and 0.5 d are aliases of the P ~6 d signal. The complex structure of this main peak is also due to our dataset's window function, causing many sharp "sidelobes". Two of these sub-peaks have similar likelihood: 5.998 d and 5.206 d. We cannot distinguish which of these is the true period. All other alias peaks and sidelobes in the periodogram have $\Delta\log L$ values more than 6.7 below the highest peak, and are therefore more than 800 times less probable (Equation 4.9). I confirmed that the main alias peaks and subpeaks seen in Figure 5.13 are reproduced by injecting a noise-free Keplerian signal at either

5.998 d or 5.206 d into a synthetic timeseries with identical sampling to our dataset (c.f. Dawson & Fabrycky 2010). Parameters of the best fit RV solutions corresponding to the highest peaks in Figure 5.13 are given in Table 5.6.

Table 5.6: Best fit Keplerian parameters for HD11231, deduced planetary parameters, and dataset properties. Columns show the results of using only HARPS data and the combined HARPS and MIKE datasets.

Parameters	HARPS	HARPS + MIKE
$\Delta \log L$	25.4	26.3
FAP [%]	6.4×10^{-5}	5.1×10^{-4}
P [d]	5.998	5.207
K [m s^{-1}]	41	38
ω [deg]	31	266
$\gamma_{\rm HARPS}$ [m s^{-1}]	-4.6	-3.9
$\gamma_{\rm MIKE}$ [m s^{-1}]	N/A	16.6
$\sigma_{\rm HARPS}$ [m s^{-1}]	17.8	18.4
$\sigma_{\rm MIKE}$ [m s^{-1}]	N/A	15.1
$M_{\rm p} \sin i$ [M$_{\rm J}$][a]	0.47	0.41
a [AU][a]	0.073	0.066
$N_{\rm obs}$	49	56
Data baseline [d]	476	4508
t_0 [BJD-2400000]	57286.78871	53254.84461

[a] assumes M_*=1.44 M$_\odot$ (Table 5.5)

There is a large RV dispersion about the best fit to our HARPS data: the residuals show peak-to-peak variability of up to 80 m s^{-1} on timescales of a few days, and the jitter value is 17.8 m s^{-1}. Searching for a second Keplerian signal in the recursive likelihood periodogram framework revealed no further signals with FAP < 1%. This is surprising: the significant RV variability remaining in the timeseries suggests prominent additional signals may be present. Stellar jitter at this level is only expected for highly active stars (see Section 5.3.4).

Testing different RV models

Forcing zero eccentricity has entirely negligible effects on signal detection and the best fit solution. The periodogram is unchanged and the fitted period and RV semi-amplitude agree within 0.0005 days and 0.1 $\mathrm{m\,s^{-1}}$ respectively. The best fit eccentricity is 0.06. For all other comparisons I forced $e = 0$.

I tested a period search including the archival MIKE dataset, allowing for independent γ_{subset} and σ_{subset} values. As seen in Figure 5.13, the likelihood periodogram remains similar with MIKE data, with changes in the peak height of prominent sidelobes. The best fit solution shifts to $P = 5.207$ d (see Table 5.6 for full parameters). RVs from MIKE and all HARPS runs consistently follow the best fit solution as seen in Figure 5.14. In Section 5.3.4 a very large

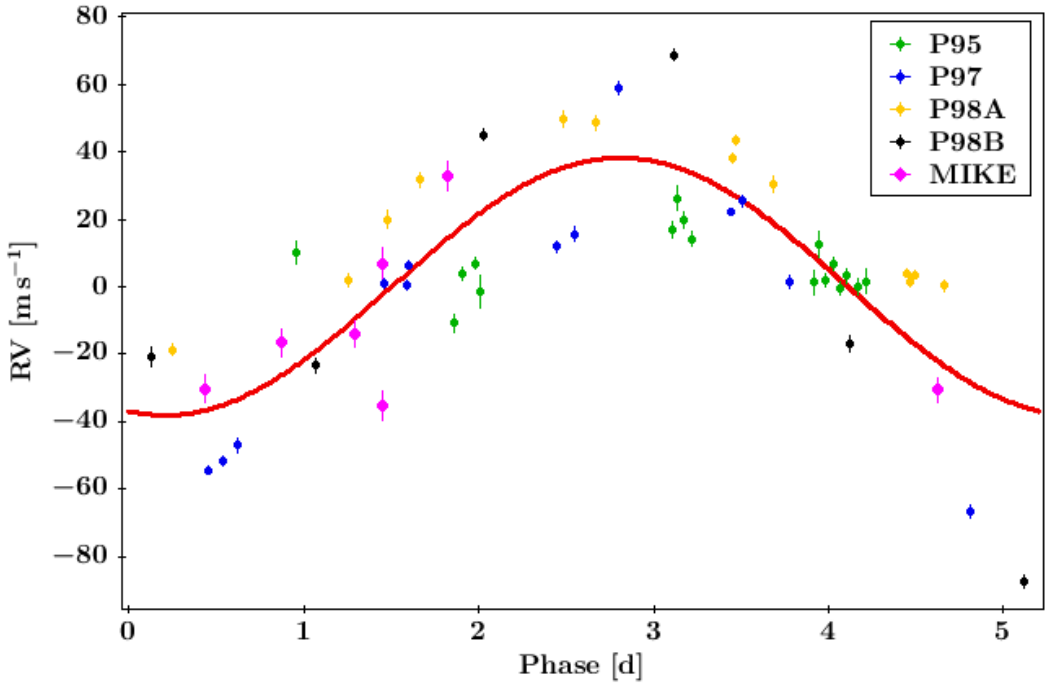

Figure 5.14: Best fit Keplerian solution for HD11231 MIKE+HARPS data. Different observing runs are colour-coded to show the phasing is relatively consistent, with significant RV scatter in each run.

(~350 $\mathrm{m\,s^{-1}}$) systematic FWHM offset between the P95-P97 and the P98 subsets will be identified. An instrumental origin can be ruled out, as this was not seen in the data of any other stars observed in these runs. The cause is therefore

a long-term systematic change in the HD11231 stellar line profile shape. This could lead to a corresponding systematic RV shift. I therefore tested the effect of allowing different $\gamma_{\rm subset}$ values for the P95-P97 and P98 subsets at the signal search level, using only HARPS data. The periodogram remained very similar to the HARPS+MIKE case, with a substantial increase in likelihood for the 5.207 d signal (Fig. 5.13). Changes in the best fit period and semi-amplitude were negligible (Table 5.7). Compared to the HARPS best fit with a single $\gamma_{\rm subset}$

Table 5.7: Best fit Keplerian parameters for three different HD11231 RV models using only HARPS data and associated $\Delta\log L$ values. Columns show the results of treating P95-P97 and P98 data as separate subsets; and of using BIS and FWHM as correlation terms.

Parameters	[P95-97 + P98]	[P95-97 + P98, ξ_i = FWHM$_i$]	[P95-98, ξ_i = BIS$_i$]
$\Delta\log L$	31.5	30.5	25.2
P [d]	5.205	5.206	5.997
K [m s^{-1}]	40	43	41
ω [deg]	11	11	30
$\gamma_{\rm HARPS}$ [m s^{-1}]	-12.9, 8.6	-13.4, 7.9	-4.5
$\sigma_{\rm HARPS}$ [m s^{-1}]	13.5, 17.7	13.1, 17.2	17.6

value, the residual RMS is reduced by 2.5 m s^{-1}. The offset between the fitted HARPS $\gamma_{\rm subset}$ values is 21.5 m s^{-1}. **In summary, none of the tests described above alter the detection of a highly significant ~6 day, K ~40 m s^{-1} period, corresponding to a companion with approximately the mass of Saturn.** The window function of our data causes the best fit solution to "jump" between the 6.0 and 5.2 day sidelobes, depending on the details of the RV model and combination of subsets used. To verify a planetary origin for this signal, purely stellar explanations must be investigated; especially given the large residual scatter.

5.3.4 A stellar origin of the RV period?

In general, phase coherence of a stellar signal would be surprising over the 1.3 year baseline of our observations (12.4 years if the archival MIKE data is added). The high amplitude and relatively short period of the signal described above might nevertheless be explained by either the rotation of a star with prominent active regions, or γ-Doradus pulsations (c.f. Table 4.1). In the following, I examine the behaviour of the HD11231 line profiles parameterised by the BIS and FWHM of the CCF, in light of the planetary and stellar signal hypotheses.

BIS and FWHM variability

Figure 5.15 illustrates that the BIS and FWHM are strongly variable on timescales similar to but distinct from the RV period.

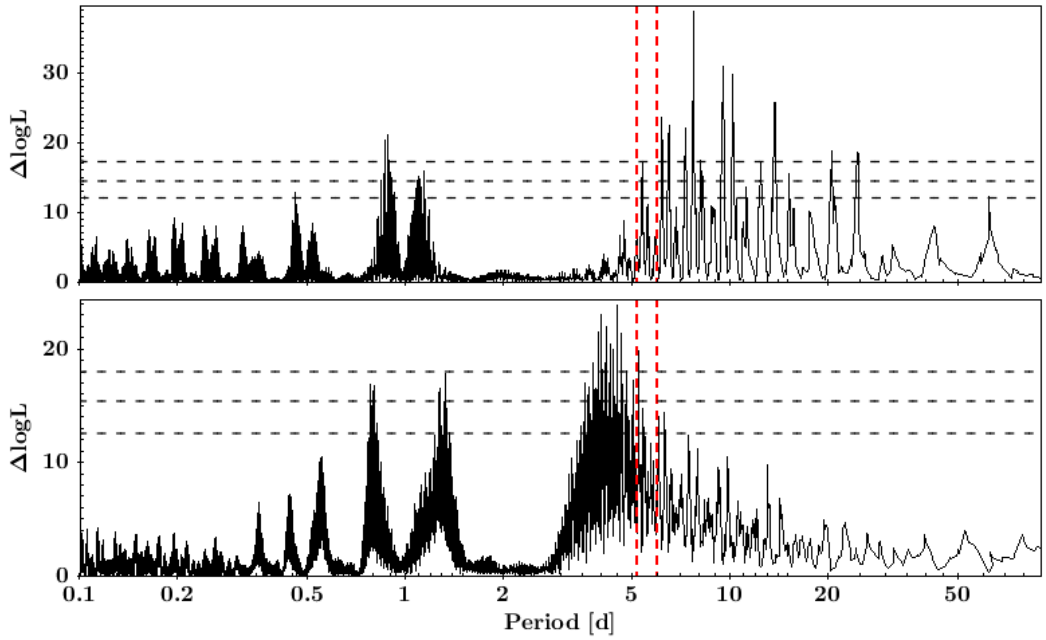

Figure 5.15: HD11231 likelihood periodograms for the BIS (top) and FWHM (bottom) of our HARPS data. In the FWHM case an offset between the P95-P97 and P98 subsets has been added at the signal search level to reveal short-period variability in addition to the long-term, high-amplitude FWHM shift (Fig. 5.16). The positions of the RV signal peaks at 5.998 d and 5.206 d are highlighted (red dashed), and the 10%, 1% and 0.1% FAP thresholds are shown.

However, as shown in Figure 5.16, there is a complete lack of correlation between RVs and the BIS & FWHM values. I confirmed this via OLS and the p-value associated with the F-test included in STATSMODELS. For the linear fits of HARPS BIS-RV and FWHM-RV data, the p-values are above 20%. The BIS and RV measurements have a very similar amplitude, while the ~500 $\mathrm{m\,s^{-1}}$ peak-to-peak FWHM variability is dominated by the offset in P98. The peak-to-peak amplitude ratio BIS/RV is 1.3, and the RMS of the FWHM and BIS is 173 and 44 $\mathrm{m\,s^{-1}}$ respectively.

This CCF parameter variability is unusual for inactive, slowly rotating FGK-type stars. In a large HARPS survey of early-type stars (Borgniet et al. 2017), only 3 out of 19 stars with similar basic parameters to HD11231 ($0.3 < B-V < 0.55$ and $v \sin i < 10$ $\mathrm{km\,s^{-1}}$) showed a BIS RMS comparable to HD11231. One of these, HD138763, is highly active with $\log(R'_{\mathrm{HK}})$= -4.4 (Borgniet et al. 2017). This star shows a clear RV-BIS correlation that is convincingly explained by the rotation of active regions (Lagrange et al. 2009; Fig. 11 in Desort et al. 2007).

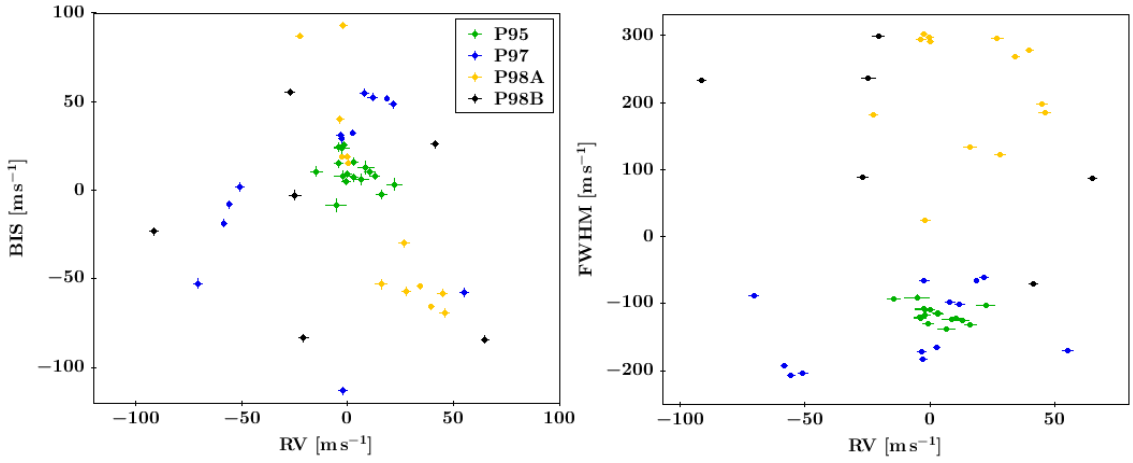

Figure 5.16: HD11231 HARPS RVs plotted against FWHM (**left panel**) & BIS (**right panel**). For each parameter the mean has been subtracted, which leads to apparent negative FWHM values.

Typical RMS values for other HARPS targets in our RV programme with a meaningful amount of data (N_{obs}>17, c.f. Section 5.1) ranged between 1-4 m s^{-1} (BIS) and 2-8 m s^{-1} (FWHM). Given the significant CCF variability, I tested RV signal recovery including an activity correlation terms with the BIS and FWHM timeseries as the $\xi_{i,\mathrm{subset}}$ values (c.f. Section 4.5). For the FWHM case, I again treated the P95-P97 and P98 RV data as separate subsets. The RV periodograms and Keplerian best fit parameters were essentially unaffected compared to the reference solutions without correlation terms (Tables 5.6 and 5.7).

Bisector shape variations

The HARPS bisectors themselves clearly show parallel shifts on the order of ± 40 m s^{-1} (Fig. 5.17) in addition to significant shape distortions (bisector "tilts"). Figure 5.17 illustrates that parallel-shifts, consistent with stellar centre-of-mass motion, clearly account for significant bisector variability in velocity space. Our Keplerian solution removes most of the parallel-shift component. This is consistent with orbital reflex motion, combined with significant residual stellar variability of the bisector.

Star spots / stellar rotation

The stellar rotation period P_* of HD11231 is unknown. Our $v \sin i$ and R_* values (Table 5.5) impose an upper limit of $P_* \lesssim 19$ days. If the rotation period were in fact 6 days, a near pole-on orientation of $i \sim 19°$would be required to match the $v \sin i$ value.

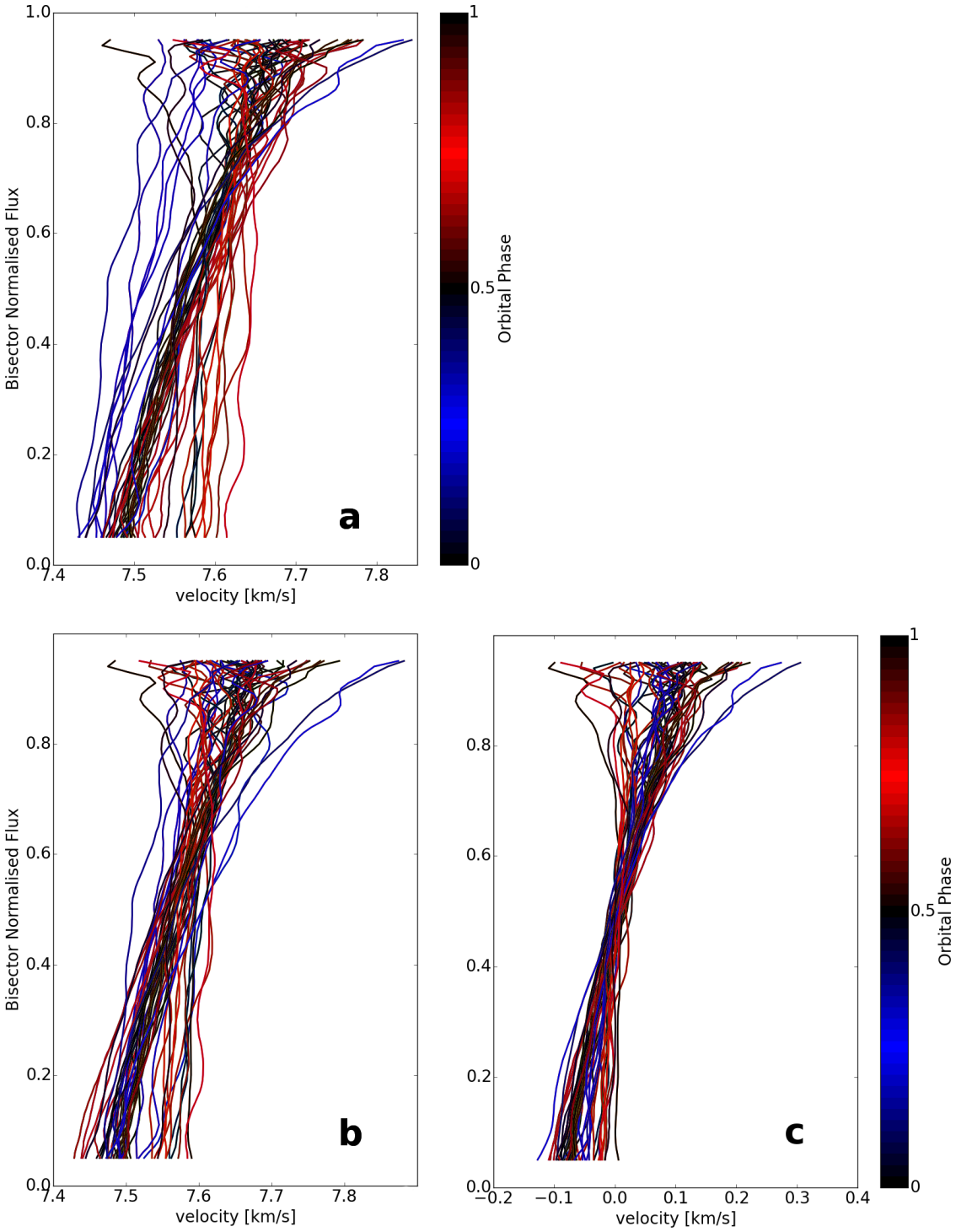

Figure 5.17: **(a)** Bisectors of HD11231 for each HARPS epoch, color-coded by phase of the best fit 5.998 day orbital solution (c.f. Table 5.6). Quadrature points of the orbit are blue and red, while conjunctions are black. **(b)** Same bisectors, parallel-shifted by the velocities of the best fit 5.998 d RV solution, evaluated at each epoch. Note that using the 5.206 d orbital solution gives very similar results. **(c)** Bisectors shifted by the measured TERRA velocity at each epoch.

There are several lines of evidence against a stellar rotation origin of the RV signal:

- Arriagada (2011) used the empirical relationships of Wright (2005a) and Isaacson & Fischer (2010) to calculate expected RV jitter values, given the relevant stellar parameters and $\log(R'_{\rm HK})$ values. For HD11231 these estimates of the stellar RV jitter contribution are 2-3 m s^{-1}. To cause stellar jitter at the 30 m s^{-1} level of our observations, a very active star with $\log(R'_{\rm HK})$ ~-4.4 would be expected (Hillenbrand et al. 2014), corresponding to an age of only ~0.3 Gyr (Mamajek & Hillenbrand 2008). This contradicts our isochronal age estimate of 2 ± 0.3 Gyr (Section 5.3.1).

- It is also implausible that interstellar absorption could lead to a putative ~0.9 dex difference between the intrinsic and measured activity level (Table 5.5). Using the online tool[1] developed in Fossati et al. (2017) I calculated this would imply an exceptionally high ISM column density of $\log(N_{\rm CaII}\,[{\rm cm}^{-2}]) \gtrsim 15.3$. Our reddening value (Table 5.5) is completely inconsistent with such a dense ISM along our line of sight to HD11231. Welsh et al. (2010) showed that $\log(N_{\rm CaII}\,[{\rm cm}^{-2}]) \lesssim 13$ is seen within 800 pc. HD11231 lies at a distance of only 134 pc.

- A clear anti-correlation between BIS and RV would be expected if this large amplitude signal were due to stellar rotation (c.f. Section 4.2 and Figure 11 in Desort et al. 2007). Figure 5.16 illustrates this is not the case. Parallel shifts of the bisectors, as seen in Figure 5.17, would not dominate if the primary source of RV variability was from rotating active regions (c.f. Queloz et al. 2001).

[1] http://geco.oeaw.ac.at/software.html

In summary, the hypothesis of a stellar rotational signal can be conclusively ruled out. Similarly, a stellar rotation signal cannot explain the 17.8 m s^{-1} RV jitter corresponding to the best fit Keplerian orbit in Section 5.3.3. At this jitter level, the expected intrinsic activity of $\log(R'_{\rm HK}) \sim$ - 4.5 (Hillenbrand et al. 2014) would imply[1] an equally implausible $\log(N_{\rm CaII}\ [{\rm cm}^{-2}]) \gtrsim 15.2$ to explain the measured $\log(R'_{\rm HK})$= -5.26. Note the column density required and the estimated intrinsic activity level are very similar to that estimated above for the initial RV RMS of 30 m s^{-1}. This is due to the logarithmic relationship between stellar RV jitter and and activity level (c.f. Figure 1.5). To explain these high column densities, both interstellar and strong circumstellar absorption contributions could be invoked speculatively. However, this cannot be reconciled with a stellar rotation origin of the RV signal either: a clear BIS-RV anti-correlation would still be observed in this scenario. Note that a contribution to the residual RV variability from stellar activity cannot be ruled out. However, I found no evidence for a correlation between the residual HARPS RVs and BIS: the p-value associated with the OLS linear best-fit is 27%. This strongly suggests that the dominant jitter component arises not from stellar rotation, leaving pulsations as the likely alternative.

γ-Doradus pulsations

γ-Doradus pulsators are A-F type main sequence stars that show multi-periodic photometric and spectroscopic variability due to non-radial gravity-mode pulsations (Kaye et al. 1999). A review of theoretical work on the pulsation mechanism can be found in Dupret (2007). Canonical pulsation periods are between 0.4 and 3 days (Kaye et al. 1999), with typical photometric amplitudes of 0.002 - 0.1 mag (Henry et al. 2011). This is distinct from the δ-Scuti pulsations of some A-F type stars, which have shorter periods down to 0.5 hours and arise from p-

mode oscillations (e.g. Breger 1979). δ-Scuti stars occupy a different instability strip in the HR diagram, bluewards of the γ-Doradus region, although there is some overlap and Hybrid behaviour (e.g. Henry & Fekel 2005). γ-Doradus pulsations are not particularly rare: estimates indicate they are present in up to ~20% of A7 to F5 type stars (Henry et al. 2011). The "classical" γ-Doradus instability strip was based on ground-based observations and has been revised with the advent of large space-based photometric surveys. These high precision measurements (Bradley et al. 2015) identified lower-amplitude (~ 0.1 mmag) pulsations and extended the instability strip to significantly cooler stars, reaching $T_{\rm eff}$ ~6100 K, i.e. spectral type F8 . The spectral line profile variability of γ-Doradus stars has been studied in detail, giving insights into stellar structure (e.g. Mathias et al. 2004, Maisonneuve et al. 2011, Davie et al. 2014). Typical RV amplitudes induced by the pulsations are 2-4 km s^{-1} (Henry et al. 2011 and references therein).

There have been few planet discoveries around pulsating stars - known pulsators are in fact avoided in planet search programmes (e.g. Lagrange et al. 2009). A notable discovery was the transiting Hot Jupiter WASP-33 orbiting a rapidly rotating A-type star (Cameron et al. 2010) showing δ-Scuti pulsations (Herrero et al. 2011). Recently short-period (~90 min), low-amplitude photometric pulsations were discovered for the eccentric HJ host HAT-P-2 (de Wit et al. 2017). These may be caused by tidal SPI and lead to the observed stellar RV jitter of 36 m s^{-1}. Note that pulations excited by tidal effects have been observed in eccentric stellar binary systems (e.g. Welsh et al. 2011 and references therein). Lagrange et al. (2009) present results from a large HARPS study of 185 A-F type stars, including several pulsators. They note that pulsating stars do not show correlations between BIS and RV. Instead, the BIS timeseries is much more variable than the RVs, with BIS/RV ratios $\gtrsim 3$. This leads to a vertical spread of points in the RV-BIS plane. As Lagrange et al. (2009) point out (c.f.

their Figure 5), the combination of pulsations and an orbital companion can lead to a "composite" distribution in the RV-BIS plane: both parameters show large-amplitude variability with no correlation. This is the behaviour seen for HD11231 in Figure 5.16, where BIS/RV = 1.3. Additional examples of this scenario can be found in Borgniet et al. (2017). Note that those works highlight pulsation-related composite behaviour with large amplitudes (1-2 km s^{-1}) of both RV and BIS, corresponding to pulsators with stellar companions.

To my knowledge, the only exoplanet host that has been explicitly identified as a likely γ-Doradus star is WASP-118. The Kepler light curve of this F6 star hosting a transiting HJ (Hay et al. 2016) exhibits low amplitude (~200 ppm), multi-periodic pulsations at periods between 1 - 2.5 days (Močnik et al. 2017). No pulsation-related analysis of WASP-118's RV and line profile variability has been done to date. Unfortunately the majority of RV data collected for WASP-118 has relatively low precision and is not publicly available (Hay et al. 2016). This hampers assessing the effect of pulsations on the RVs and exact comparisons with HD11231. Intriguingly, the BIS-RV plot presented in Hay et al. (2016) does show a composite behaviour comparable to my Figure 5.16. The WASP-118 data also show no correlation, and the peak-to-peak variability of both BIS and RV reaches 300 m s^{-1}. Hay et al. (2016) do not quote the stellar RV jitter and RMS of their WASP-118 data, but note that the orbital reflex amplitude is K= 55 m s^{-1}.

Galland et al. (2006) provide an in-depth discussion of the RV discovery of a 25 M_J Brown Dwarf orbiting the γ-Doradus A9V star HD180777. In this case, peak-to-peak RV variations of 1.7 km s^{-1} due to pulsations are left in the time series after subtracting the best fit Keplerian signal. The residual RMS is 0.4 km s^{-1} and the amplitude ratio BIS/RV reaches 1.2. Galland et al. (2006) show that the Brown Dwarf signal corresponds to parallel shifts of the CCF, i.e. centre-of-mass motion. The HD11231 system could be seen as a "scaled

down" version of the HD180777 case, with a lower mass star and companion, combined with lower amplitude pulsations. Additional precision RVs for the WASP-118 system would clarify whether this is a more direct analogue to HD11231.

5.3.5 Summary: HD11231

In summary, the BIS and RV variability of HD11231 is consistent with a giant planet orbiting a pulsating star. For an F5V star such as HD11231, low amplitude γ-Doradus pulsations are plausible. Note that in the context of the work in Grassitelli et al. (2015), our measured v_{mac} value and HD11231's position on the HR-diagram are also consistent with a γ-Doradus star. The effect of the pulsations on our BIS and FWHM data appears to fall in the period ranges of 9-10 and 4-5 days respectively; suprisingly large compared to the canonical 0.3 - 3 day γ-Doradus periods. An SPI origin of the HD11231 pulsations should be explored in future work. No clear pulsation periods are detected in our RV measurements after the planetary signal is fit. The pulsation-induced RV variability seen in our data is at a low level compared to well-studied γ-Doradus stars, and the detection of corresponding photometric pulsations could be challenging. There is no convincing evidence in our RV data for additional planets in the HD11231 system. The presence of pulsations and associated stellar line-profile distortions greatly complicate the search for any (lower RV amplitude) companions.

5.4 A compact multi-planet system around HD38677

5.4.1 Stellar properties

HD38677 is an 8th magnitude F8V star in the Orion constellation. This target was selected from our P13 sample due to its median activity level of $\log(R'_{HK})$ = -5.16 from the six observations of Arriagada (2011). As for HD11231, we had access to MIKE RVs from the Magellan Planet Search. In this case, only 4 measurements over a baseline of 120 days are available. These velocities did not show significant variability in the context of MIKE's ~5 m s^{-1} stability, but this is too small a dataset to be particularly meaningful.

A precision analysis of the stellar parameters is included in the re-analysis of the Geneva-Copenhagen survey (Casagrande et al. 2011), hereafter C11. This study makes use of the infrared flux method and Stromgren photometry, recovering the stellar bolometric flux, effective temperature, gravity, iron abundance [Fe/H], overall metal content [M/H] and α-element content [α/Fe]. A Baysian approach to isochrone fitting is then used to estimate stellar mass and age. The inputs for this are the absolute V magnitude from Hipparcos, [M/H], $T_{\rm eff}$ and a Salpeter mass function (Salpeter 1955). As done throughout C11, we adopt their expectation values derived with BASTI isochrones for M_* and stellar age. The PADOVA isochrone results also provided by C11 for HD38677 are in close agreement with these values. Stellar parameters are listed in Table 5.8.

5.4.2 Observations

We began observations of this target in P96, collecting a high cadence dataset over 5 nights in excellent observing conditions. Figure 5.18 illustrates that the RVs primarily varied on timescales of > 1 day, by ~4 m s^{-1} peak-to-peak; i.e. at a much greater level than the interleaved observations of the RV standard

Table 5.8: HD38677 stellar parameters.

Parameter	value	Reference
Spectral Type	F8V	Hipparcos[a]
Parallax [mas]	16.03 ± 0.25	TGAS[b]
Distance [pc]	62 ± 1	TGAS
V [mag]	7.98	Hipparcos
$B-V$ [mag]	0.581 ± 0.013	Hipparcos
$\log(R'_{\rm HK})$	-5.16	Arriagada 2011
$T_{\rm eff}$ [K]	6176 ± 74	C11
[Fe/H]	0.12	C11
[M/H]	0.12	C11
[α/Fe]	-0.05	C11
$\log g$ [cm s^{-2}]	4.32	C11
M_* [M$_\odot$]	$1.17^{+0.04}_{-0.03}$	C11
L_* [L$_\odot$]	2.0	C11
Age [Gyr]	$2.7^{+1.3}_{-1.5}$	C11

[a]XHIP compilation; Anderson & Francis 2012
[b]Gaia Collaboration et al. 2016

GJ191. See Section 5.2.2 for details on GJ191. Note the RMS of the P96 HD38677 and GJ191 datasets was 2.3 and 0.8 m s^{-1} respectively (c.f. Figure 5.10), with median internal uncertainties of 0.7 and 0.6 m s^{-1}.

Likelihood periodogram analysis of the P96 data showed a highly significant (FAP = 2.3 $\times 10^{-9}$) signal at a period of 4.6 days. A signal with P = 2.9 d had a near-equal $\Delta \log L$ value, and provided an equally good fit to the data as illustrated in Figure 5.18. Clearly additional data was needed to distinguish these solutions and to confirm a truly periodic signal. The P96 data is a good example for the problems of detecting formally high-significance signals with periods of a few days in an observing run with a baseline of 5 nights. Note that the 1 night gap seen in Figure 5.18 arose from ESO run scheduling rather than bad weather. Without this gap, the degeneracy between the two solutions shown would be lifted. As already illustrated in Figure 4.4 of Section 4.3.1, the CTI correction has a negligible effect on these findings.

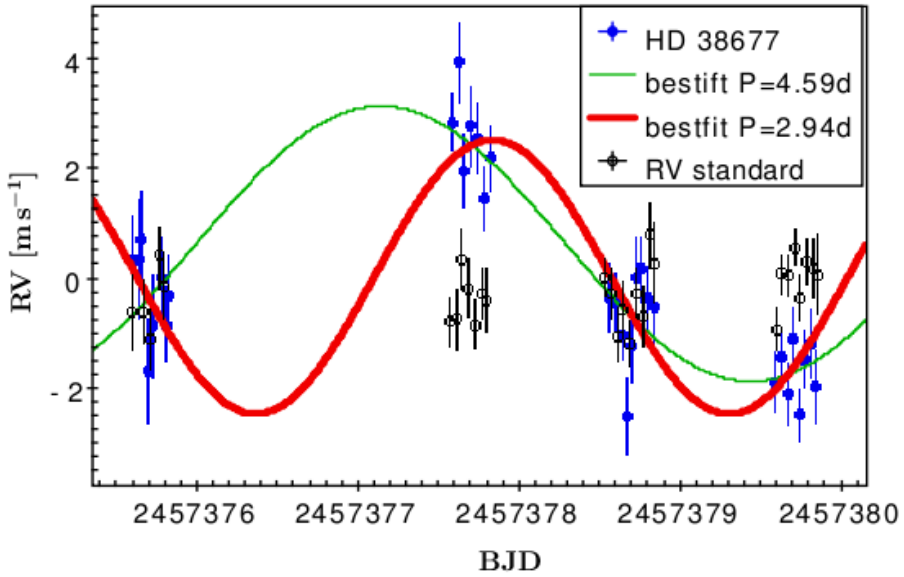

Figure 5.18: HARPS P96 timeseries of HD38677 and the RV standard GJ191, along with the two alternative best fit signals from likelihood periodogram analysis.

High cadence observations of HD38677 continued in our P98 runs, along with 5 low-cadence points provided by Co-I James Jenkins via a time-share agreement. In total, 113 observations were collected (Figure 5.19): 32 points in P96; 35 points in P98(A); 5 points in P98(J) and 41 points in P98(B). The entire dataset's RV RMS was 4.0 m s^{-1}. Exposure times for all but two observations were 900 seconds. Variable extinction and seeing led to a SNR range of 70-250 (median of 175) and σ_{RV} values between 0.5 and 2.1 m s^{-1} (median of 0.8 m s^{-1}).

5.4.3 RV periodicity

As illustrated by Figure 5.19, the RV timeseries of HD38677 is dominated by 8 $m\,s^{-1}$ peak-to-peak variability on timescales of a few days (seen in all sub-runs), combined with an apparent offset of ~7 $m\,s^{-1}$ over the 342 day gap between the P96 and P98 runs. A comparison of Figures 5.19 and 5.18 highlights the more complicated RV variability revealed by P98 data compared to the single-planet best fits calculated from P96 data only.

I carried out an initial recursive period-search using all HARPS data as a single subset and forcing circular orbits throughout. The successive likelihood periodograms corresponding to this are shown in Figure 5.20, along with the dataset's window function. Recall the latter is the Fourier transform of a timeseries' sampling, and shows the alias periods of a signal with a frequency of zero (infinite period). The alias patterns seen in the likelihood periodograms of our timeseries are a direct consequence of the shape of this window function. As described in the following, I found three distinct, highly significant signals in the RV timeseries, with false alarm probabilities well below 0.1%. The 3-planet global best fit parameters are listed in Table 5.9 and the corresponding phased radial velocities are shown in Figure 5.21.

- As seen in Figure 5.20, an initial signal with a period around 20 days is detected with extremely high significance (FAP = 2.5 $\times 10^{-33}$). The distinct sets of alias peaks of this signal (at ~13 d; 1 d; etc.) have $\Delta \log L$ values smaller by at least 18.2, and are therefore 8×10^7 times less probable (Equation 4.9). While the highest likelihood solution is found at 20.9 days, the "sidelobe" at 19.8 days is only 16 times less probable.

- The search for a second signal led to a best fit period at ~10 days with a false alarm probability of 1.5 $\times 10^{-9}$. The most likely 9.97 day solution and a sidelobe peak at 10.25 days have near-equal likelihood. The closest distinct alias solution at 6.57 days is ~80 times less probable.

- Finally, I recovered a third signal at 2.88 days with FAP = 1.4 $\times 10^{-5}$. As illustrated in Figure 5.20, alias peaks at 1.69 and 2.47 days have very similar $\Delta \log L$ values. It is not clear which of the three is the true period as they have probabilities within a factor of two of each other.

- The residual radial velocity RMS of the three planet solution is 1.2 m s^{-1} and no further signals with a false alarm probability below 10% were found.

- This best fit corresponds to 2 Neptune mass planets in warm orbits, and a hot, short-period super-Earth. Due to the alias ambiguity for the third signal, a solution with a significantly different period (1.69 d) cannot be ruled out. Forcing a three planet best fit with the third signal at 1.69 days results in a very similar overall solution: Two Neptune mass planets with the same periods as before, and an inner planet with 2.5 $M_\oplus$.

Table 5.9: Best fit Keplerian parameters for 3-planet HD38677 solution, deduced planetary parameters, and properties of the dataset.

Parameters	HD38677 b	HD38677 c	HD38677 d
$\Delta \log L$	87.7	31.2	21.9
FAP [%]	2.5 $\times 10^{-31}$	1.5 $\times 10^{-7}$	1.4 $\times 10^{-3}$
P [d]	20.94	9.97	2.88
K [m s^{-1}]	3.7	3.4	1.2
ω [deg]	316	309	82
$M_p \sin i$ [$M_\oplus$][a]	18	13	3
a [AU][a]	0.158	0.096	0.042
γ_{HARPS} [m s^{-1}]	-0.3		
σ_{HARPS} [m s^{-1}]	0.7		
N_{obs}	113		
Data baseline [d]	387		
t_0 [BJD-2400000]	57375.58981		

[a] assumes M_*=1.17 $M_\odot$ (Table 5.8)

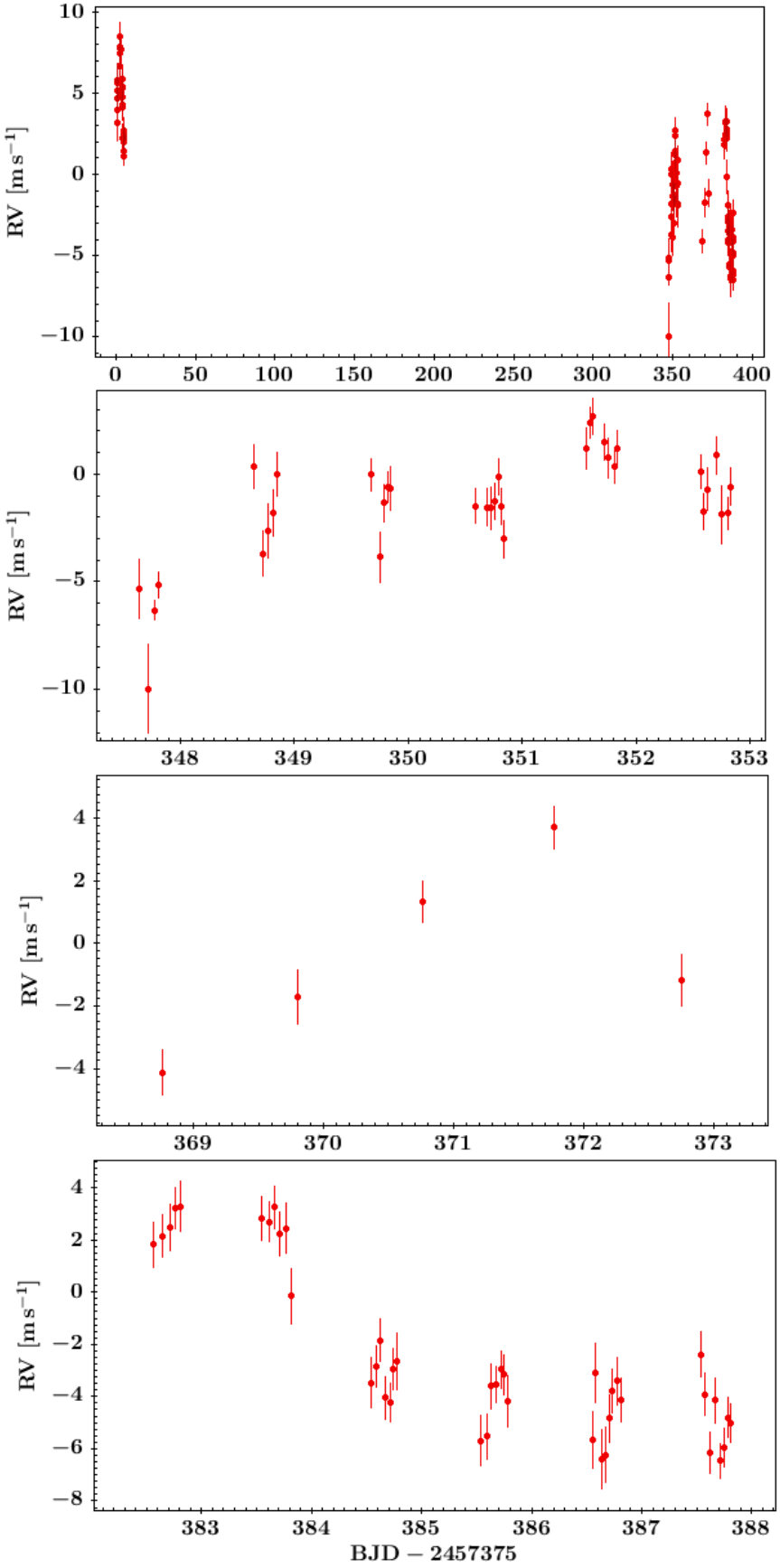

Figure 5.19: Our entire HARPS timeseries of HD38677 (top panel), followed by close-up views of the P98(A), P98(J) and P98(B) runs, which span 5-6 consecutive nights. For a close-up of P96, see Figure 5.18.

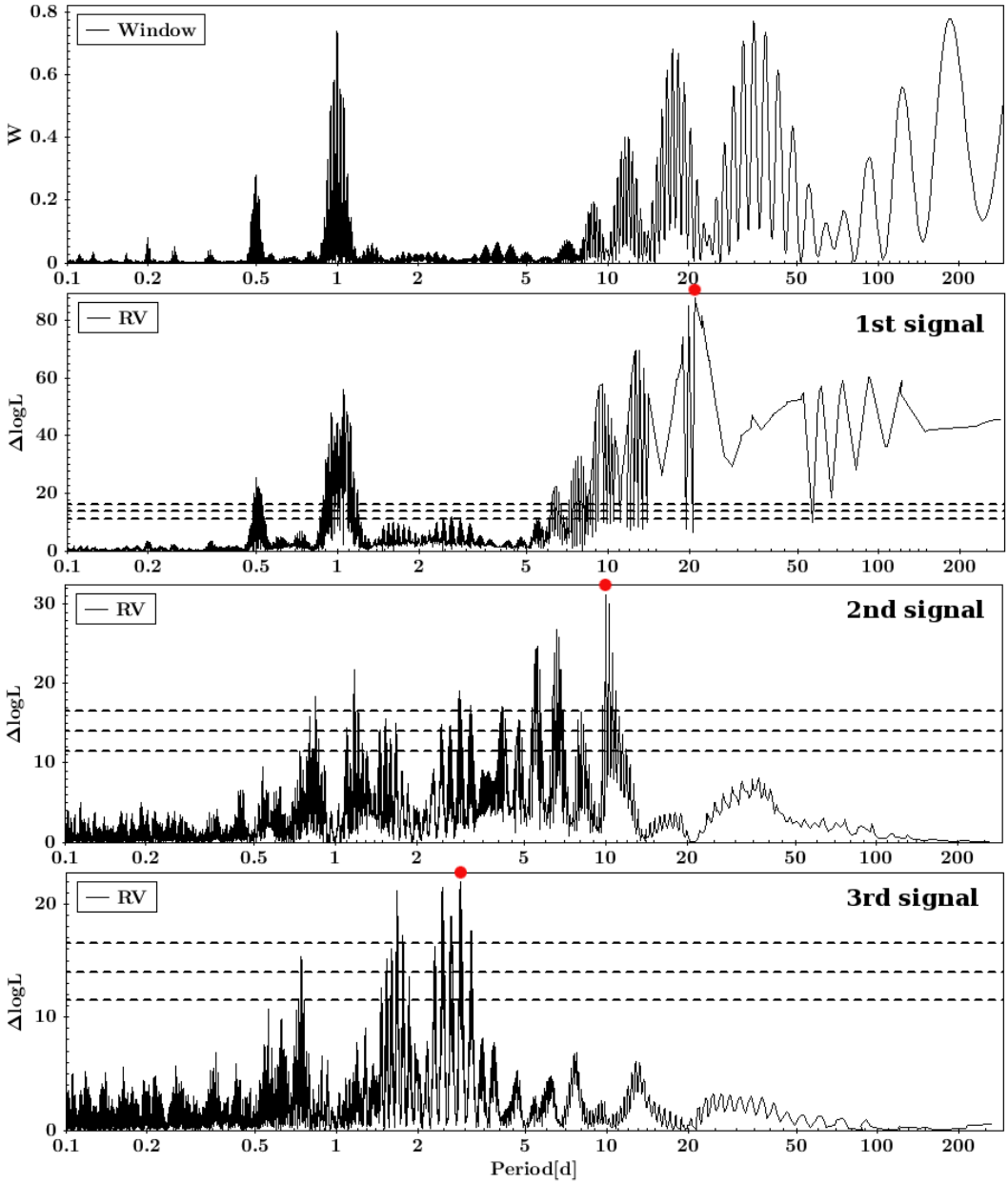

Figure 5.20: HD38677 RV likelihood periodograms and the dataset's window function. The three RV periodograms generated during the recursive period search are plotted and the most probable periods are highlighted. 10%, 1% and 0.1% FAP thresholds are represented by dashed lines.

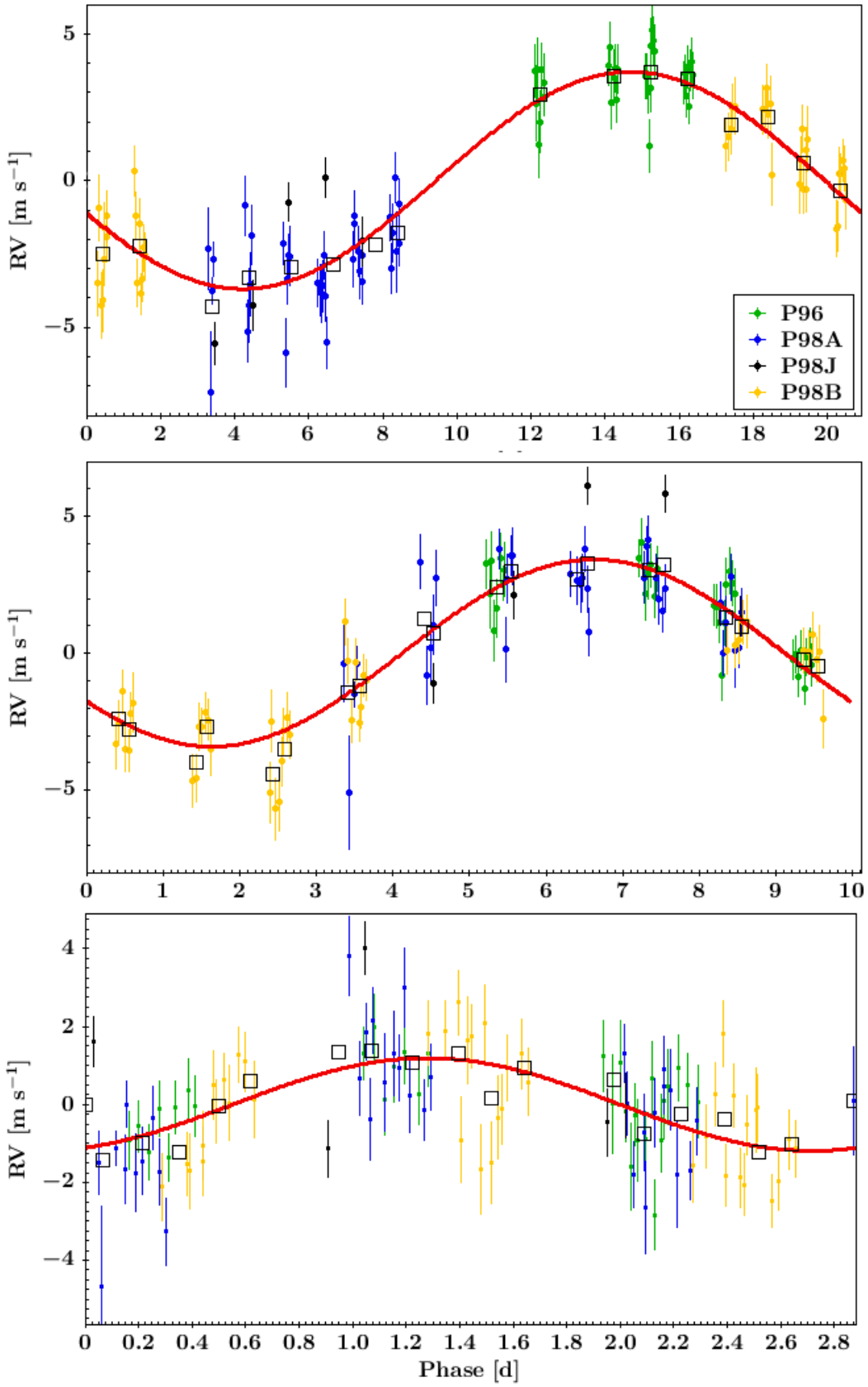

Figure 5.21: Phased best fit Keplerian solutions for HD38677 RVs. For each panel, the other two Keplerian signals detected have been removed. Different observing runs are colour-coded, and binned observations (squares) are shown for illustration.

5.4.4 Stellar signals

In contrast to HD11231, the HD38677 observations show no prominent line profile distortions on timescales of a few days or less. This is evident from the HD38677 bisector shape over time, which is dominated by parallel shifts in velocity (Figure 5.22), and from the flat BIS and FHWM timeseries observed (Figure 5.23). The RMS values of those parameters are 2 and 3 m s^{-1} respectively.

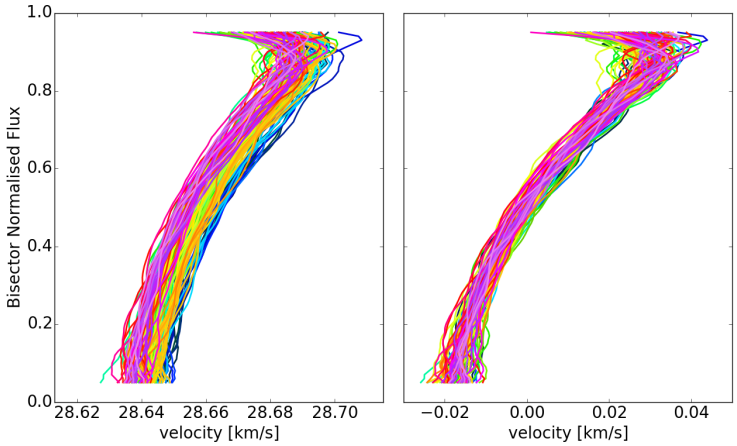

Figure 5.22: **(Left:)** Bisectors of HD38677 at each HARPS epoch excluding the observation with lowest SNR. **(Right:)** Same bisectors, parallel-shifted by the measured TERRA radial velocity. Colour-coding is arbitrary and was done for better visualisation.

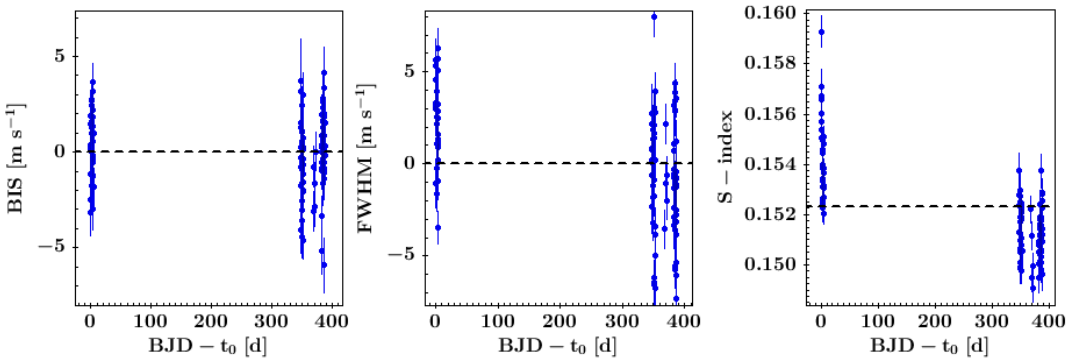

Figure 5.23: BIS, FWHM, and S-index timeseries of HD38677 covering our P96-98 observations. The mean value of each parameter is highlighted (dashed), and has been subtracted in the case of BIS and FWHM. Both the S-index and the FWHM show an apparent offset over the 342 day gap between the P96 and P98 runs.

To identify any periodic low-amplitude stellar line profile changes and stellar activity variability, I calculated the likelihood periodograms for BIS, FWHM and S-index measured from our HARPS data (Figure 5.24). S-values were generated from TERRA, and a single, highly discrepant measurement recorded in P98(A) was excluded from all analysis in this section. This 15 σ outlier above the median activity level of the dataset occurred for the observation with by far the lowest SNR. The corresponding RV, FHWM and BIS values were not significant outliers relative to their respective timeseries. The activity value likely suffers a spurious systematic offset arising from ThAr contamination within the S-index core bandpasses for this low SNR stellar spectrum (c.f. Sections 2.2 and 4.1).

Figure 5.24 illustrates that there are no significant periodic signals present in the BIS timeseries. In contrast, the S-index periodogram shows extremely significant signals well below the 0.1% FAP threshold. However, the pattern of alias peaks clearly matches that of the window function. This is consistent with a linear trend, or a poorly constrained signal with a period much longer than the baseline of observations, as already suggested by the timeseries in Figure 5.23. Adding a linear trend at the signal search level removes the putative signals with periods longer than 20 days, leaving signals with less than 0.1% FAP close to 1 and 0.5 days, and at 17.6 and 18.5 days. The FWHM timeseries shows low-significance signals (clearing the 10% false alarm probability threshold) that for the most part match those of the S-index. Again, adding a linear trend to the likelihood model removes longer period signals. In summary, there are no peaks with less than 0.1% FAP directly matching the three RV signals. However, the S-index shows signals around 1% FAP between approximately 7 - 20 days (close to 2 of the RV periods) when a linear trend is included in the fit.

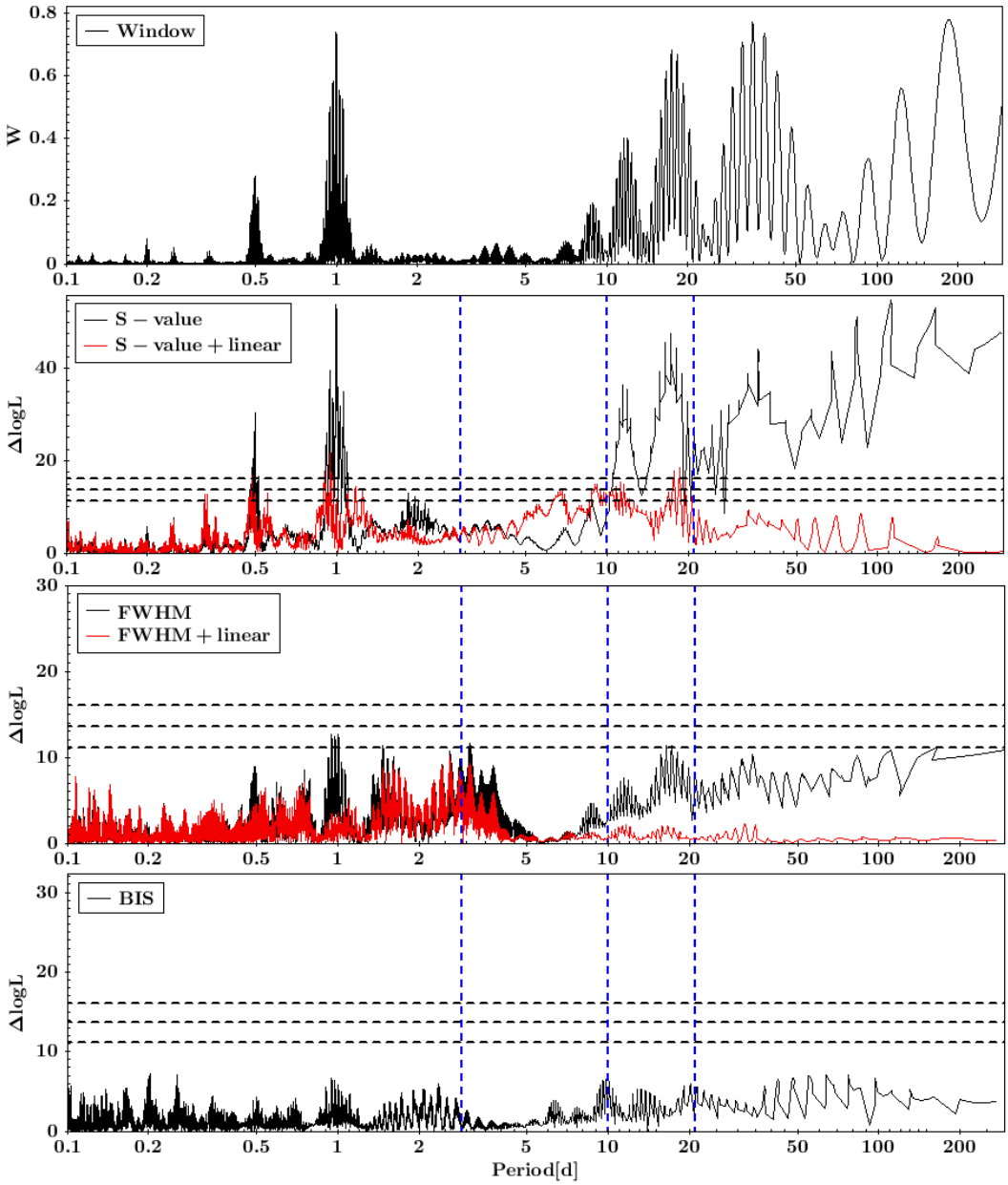

Figure 5.24: HD38677 likelihood periodograms for the S-index, FWHM and BIS. Periodograms including a linear trend in the likelihood model are shown in red. The three main RV peaks at 20.94 d, 9.97 d and 2.88 d are highlighted (blue dashed), along with the 10%, 1% and 0.1% FAP thresholds. The window function is plotted again to illustrate the alias pattern expected from long-term trends.

As an additional check, I tested the significance of linear best fit correlations between RV and the proxies tracing potential stellar signals (FWHM, S-index and BIS). Results are shown in Table 5.10 and Figure 5.25. As done in Section 5.2, I used OLS and the p-value associated with the F-test included in STATSMODELS.

The BIS does not correlate with the RVs regardless of which data subset is

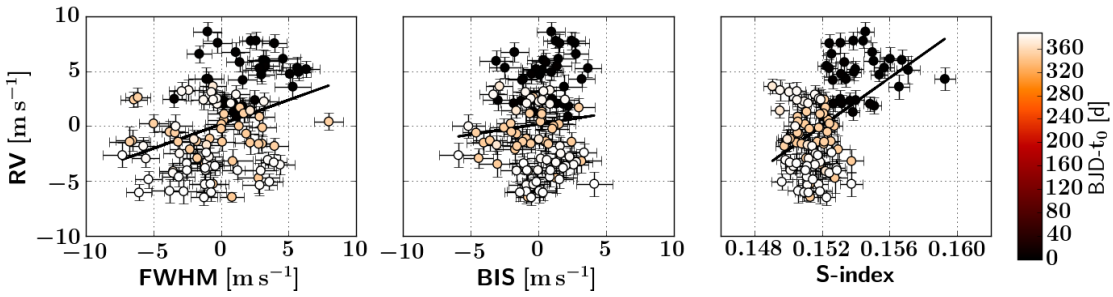

Figure 5.25: HD38677 correlations between RV and stellar signal proxies, with best fit results from OLS (Table 5.10). There are clear offsets occurring over the ~1 year gap between P96 (black) and P98 (orange-white).

Table 5.10: Results of OLS linear fits to RV and stellar signal proxies.

Data	FWHM-RV slope	p-value	BIS-RV slope	p-value	S-index-RV slope [m s^{-1}]	p-value
All	0.4 ± 0.1	1.3×10^{-4}	0.2 ± 0.2	0.32	$(1.1 \pm 0.2)\times10^{3}$	1.7×10^{-8}
P96	0.2 ± 0.2	0.30	0.2 ± 0.2	0.40	$(0.1 \pm 0.2)\times10^{3}$	0.64
P98	0.1 ± 0.1	0.53	-0.2 ± 0.2	0.29	$(-0.6 \pm 0.3)\times10^{3}$	0.07

chosen. Figure 5.25 illustrates that the RV variability significantly exceeds that of the BIS, with a peak-to-peak BIS/RV ratio of 0.7. This is evidence against stellar rotation or pulsations driving the RV variability, as explained in detail for the case of HD11231 in Section 5.3.4 (c.f. Desort et al. 2007 and Lagrange et al. 2009). There are highly significant correlations between the radial velocities and both the FWHM and S-values of the full HD38677 timeseries (top row of Table 5.10). Figure 5.25 illustrates how this is driven by the offsets seen for these 3 parameters between the P96 and P98 subsets. I confirmed that no correlations are seen at high significance within either subset in isolation (Table 5.10): the smallest p-value is 7% for the RV - S-index correlation in P98. The correlations and offsets observed on a timescale of 1 year could be driven

by a long-term stellar activity cycle: Lovis et al. (2011) showed that RV semi-amplitudes typically between 0 - 3 m s^{-1} and up to 10 m s^{-1} can be induced (see Section 4.2). The presence of RV correlations with FWHM and S-values in the overall dataset and the activity proxy periodicity being somewhat similar to that of the RVs makes further checks mandatory.

5.4.5 Testing different RV models

I carried out recursive likelihood periodogram analysis for the HD38677 RVs as in Section 5.4.3, while incorporating either the FWHM or the S-values into the likelihood model via activity correlation terms (see Section 4.5).

- The global best fit including an RV-FWHM correlation results in a near-identical three planet solution compared to the original given in Table 5.9. All orbital periods agree within 0.01 days and the RV semi-amplitudes are consistent at the 0.1 m s^{-1} level. The $\Delta\log L$ values of the second and third Keplerian signals agree within 0.1 of those in the original best fit, while that of the first signal is reduced by 7.2, corresponding to a reduced FAP that still reaches 4.9×10^{-30}. Overall the likelihood periodograms obtained are very close matches to those in Figure 5.20.

- Using the S-index as the activity correlation term also leads to very similar periodograms. The three planet best fit is slightly altered: the 19.8 day side-lobe of the first signal reaches the highest likelihood, again with a reduced $\Delta\log L$ value compared to the original that nevertheless remains well above the 0.1% FAP threshold. The second signal is detected at exactly the same period and likelihood as in the original. Finally, the best fit's third signal has a period of 2.47 d, corresponding to one of the sidelobes of the original 2.88 d detection. Its $\Delta\log L$ value increases marginally. All RV semi-amplitudes and $M_{\rm p}\ \sin i$ values match those of the solution without a correlation term at the $\leq$10% level.

Since the FWHM, S-values and the RVs showed potential offsets between P96 and P98 data, I also tested a likelihood model treating these as independent subsets, i.e. allowing an RV offset and different jitter values at the signal search level. In the corresponding recursive periodogram search, two highly significant signals at ~10 and 3 days are recovered (Table 5.11). The 20 day signal is absent, and no further signals with FAP < 0.1% are found. The parameters of the resulting 2 planet best fit are very similar to the two shorter period Keplerian signals identified in the original period search (Table 5.9). The periods, semi-amplitudes and ω values agree at the ≤15% level, except for the ω value of the ~3 day signal which only agrees to within 25%. The difference between the fitted P96 and P98 γ values amounts to 5.6 m s^{-1}, and the residual RV RMS is 1.2 m s^{-1}, as was the case for the original 3 planet best fit without an offset.

Table 5.11: Best fit Keplerian parameters for 2-planet HD38677 solution obtained when treating P96 and P98 data as independent subsets. The deduced planetary parameters and properties of the dataset are also shown.

Parameters	HD38677 b	HD38677 c
$\Delta \log L$	64.4	28.5
FAP [%]	3.8×10^{-23}	2.9×10^{-8}
P [d]	10.28	3.15
K [m s^{-1}]	3.1	1.4
ω [deg]	312	109
$M_p \sin i$ [$M_\oplus$][a]	12	4
a [AU][a]	0.098	0.045
γ_{P96} [m s^{-1}]	3.4	
σ_{P96} [m s^{-1}]	0.0	
γ_{P98} [m s^{-1}]	-2.2	
σ_{P98} [m s^{-1}]	1.0	
N_{obs}	32+81	
Data baseline [d]	387	
t_0 [BJD-2400000]	57375.58981	

[a] assumes M_*=1.17 $M_\odot$ (Table 5.8)

5.4.6 Summary: HD38677

I identified three highly significant Keplerian signals in the large HD38677 RV dataset with periods of approximately 3, 10 and 20 days, corresponding to companions with minimum masses of ~18, 13 and 3 $M_{\oplus}$. I found no evidence that the two shorter period signals have a stellar origin. There is no strong evidence for a purely stellar origin of the ~20 day signal, as it is still recovered with a FAP far below 0.1% in global likelihood optimisation including RV correlations with FWHM and S-index. However, there is cause for concern: the signal's likelihood is reduced when adding these correlation terms, and it vanishes when including separate γ_{P96} and γ_{P98} values in the RV model. Additional (low-cadence) data covering a longer baseline should clarify the reality of this signal, and whether it has a planetary or stellar origin. Further observations could also resolve the ambiguity between the ~3 day signal and its alias at ~1.7 days. The current data show that in either case this signal is due to a super-Earth companion with ~ 3 $M_{\oplus}$.

Our observations are consistent with a compact planetary system made up of either 2 Neptune mass planets and a short-period super-Earth within 0.16 AU, or one Neptune and super-Earth within 0.1 AU. In the latter case significant long-term RV variability was removed by adding an offset to the best fit model, and still requires a physical explanation. Further observations may reveal a long period companion different from the ~20 day planet candidate, or show that variability on this timescale is in fact driven by stellar activity.

5.5 Summary: RV results

In this chapter I provided the first assessment of the chromatic effect for the SOPHIE spectrograph, and demonstrated a large impact on FWHM and RV values measured with the cross-correlation approach. Due to the focus on short-period planets, our targets were observed at high cadence over a significant fraction of each night, i.e. sampling a significant airmass range. I showed that under these circumstances, the chromatic effect can lead to intra-night FWHM and DRS RV variations of 150 and 15 $\mathrm{m\,s^{-1}}$ respectively for SOPHIE data. The nominal radial precision of SOPHIE can be degraded by a factor of ~5 if no corrections are applied. In contrast, I found no significant evidence for an impact on SOPHIE radial velocities obtained via template-matching. In this case, the expected 1-2 $\mathrm{m\,s^{-1}}$ precision was achieved for high cadence observations taken over 5 consecutive nights in the case of an RV standard and for a science target with no evidence for Keplerian signals.

In the case of HARPS observations, RVs obtained from both the cross-correlation and the template-matching approach were not affected by the chromatic effect. The FWHM values of the FG-type science targets in our program also showed no evidence for this systematic effect, in contrast to the FWHM of an M-dwarf RV standard. This demonstrated the importance of flux-normalisation in CCF computation. I showed that for our observations, RV timeseries obtained via TERRA displayed RMS reductions between 0.2 and 0.5 $\mathrm{m\,s^{-1}}$ compared to their DRS counterparts. The former were therefore used for the remainder of the analysis in this chapter. I demonstrated that the high cadence observations over consecutive nights used in our program are compatible with HARPS RV stability below 1 $\mathrm{m\,s^{-1}}$.

I presented the discovery of two planetary systems around the bright, nearby F-type stars HD11231 and HD38677, which were targeted due to their anomalously low $\log(R'_{\mathrm{HK}})$ values. HD11231 data required careful analysis of

prominent stellar line profile distortions. I concluded that the observations are consistent with the presence of low amplitude γ-Doradus pulsations and a Saturn mass companion with a period of approximately 6 days. This is only the second reported detection of a γ-Doradus host with a companion in the planetary mass domain. Finally, I discovered a compact multi-planet system around HD38677 consisting of 2-3 short period planets with masses in the super-Earth to Neptune range.

Details on collaborative contributions

The observing proposals of the RV programme described in this Chapter (the Dispersed Matter Planet Project) were a collaborative effort led by my supervisor C.A. Haswell, who wrote the science justifications. My input was target selection and prioritisation, based on my analysis of both archival data and RV data from relevant runs listed in Section 5.1. I led the technical case and detailed design of the observing strategy for the proposals, with contributions from C.A. Haswell, J.R. Barnes and G. Anglada-Escude. The RV observations were carried out by myself, local telescope operators, J.R. Barnes, C.A. Haswell, J. Cooper, J. Doherty, G.D. Smith and C. McCune (see Table 5.1). L. Fossati performed the entire stellar parameter analysis of HD11231 summarised in Section 5.3.1. Detailed discussions with Z.M. Berdinas and G. Anglada-Escude provided extremely valuable input for the analyis of RVs and the chromatic effect. P. Arriagada provided the archival MIKE activity and RV timeseries. G. Anglada-Escude provided the algorithms that perform recursive maximisation of the global likelihood and output likelihood ratio periodogram data.

Chapter 6

Conclusions

"Maybe it's a little early... Maybe the time is not quite yet... But those other worlds - promising untold opportunities - beckon. Silently they orbit, waiting."

-Carl Sagan, 1994

6.1 Summary of the work in this thesis

Activity measurements of known short-period planet hosts

Note that parts of this Section were published in Staab et al. (2017)

In Chapter 3, I calibrated the Robert Stobie Spectrograph (RSS) to the Mount Wilson system of stellar activity measurements. Since this instrument is attached to the 10 m class Southern African Large Telescope (SALT), efficient and precise $\log(R'_{\rm HK})$ measurements for relatively faint transiting planet hosts are enabled via its medium resolution longslit spectroscopy mode. I verified that the data reduction and the observing mode selected are reliable by showing that consecutive calibrator star activity measurements agreed at the 1% level. As an initial science application, I measured $\log(R'_{\rm HK})$ values of the Hot Jupiter hosts WASP-43, WASP-51/HAT-P-30, WASP-72 and WASP-103, as also presented in Staab et al. (2017).

Reviewing the planet host compilation in Figueira et al. (2014), I found that a quarter of the host stars of transiting, short period exoplanets with activity data exhibit anomalously low values of $\log(R'_{\rm HK})$ below the main sequence basal limit (Section 3.4). I listed these 22 hosts (a three-fold increase over those reported in Fossati et al. 2013) and estimated the extent to which absorption in the interstellar medium could contribute to apparent activity depression. In addition to Hot Jupiter hosts, including WASP-72, this group contains two low-mass multi-planet systems. I highlighted that many of these objects below the MS basal limit also have unusally low activity levels relative to more evolved stars.

WASP-43 has an exceptionally high $\log(R'_{\rm HK})$ value, both in absolute terms and relative to its stellar age, rotation period and X-ray emission (Section 3.4.2).

WASP-103 displays marginally higher activity than expected from its isochronal age, but this conclusion should be re-assessed with improved stellar parameters. Both of these HJ systems are predicted to show significant SPI, which may be the dominant mechanism affecting their $\log(R'_{HK})$ values. The activity level of WASP-51 is anomalously low for its age, but remains above the basal level. My estimates showed that absorption in the ISM is likely insufficient to explain the depressed $\log(R'_{HK})$ values of WASP-51 and WASP-72, leaving circumstellar absorption due to planetary mass-loss as a plausible hypothesis.

Activity database study

I cross-matched an archival compilation of the minimum and maximum activity levels of thousands of stars with Hipparcos data and separated the entries into main sequence and evolved subsets (Chapter 2). 39 out of 2716 objects in my MS sample showed maximum activity values below the basal limit. In addition, I found 6 young stars in the Pleiades and Hyades clusters with anomalously low mean activity levels relative to the cluster distributions. A subset of these 45 objects became the targets of a highly focussed search for short-period, mass-losing planetary companions. This search tests the hypothesis that planetary mass loss is the cause of the observed apparent activity depression.

Using the archival activity compilation I found dramatic, transient drops in several S-value timeseries at the ~30 - 90% level and largely attributed these to spurious systematic effects (Section 2.2). I showed that the Ca II transit detection reported in the literature for the USP 55 Cnc e corresponds to a clear, ~1% apparent activity depression in-transit. I searched for other transit signals in the activity timeseries of known short period planets. The Hot Jupiter HD 189733 shows an apparent Ca II transit, but further analysis suggested that a stellar origin of this signal is more likely (Barnes et al. 2016).

Corrections and systematic effects for RV measurements

Chapter 4 details the instruments and observing strategy used for the high-cadence RV planet search for short-period, mass-losing exoplanets. With simplified assumptions, detection limits were predicted to reach 2 $M_\oplus$ planets over five nights of observations for the shortest (sub-day) periods (Section 4.4). For both the SOPHIE and HARPS instruments, I used archival RV standard data to empirically derive new corrections for the charge transfer inefficiency effect over the parameter space of interest for our RV programme (Sections 4.3.1 and 4.3.2). The correction has a negligible impact on HARPS radial velocities, but is important for SOPHIE. I applied several additional corrections to the RV output of the latter instrument, resulting in significant reductions in RV scatter (Section 4.3.2). The first assessment of the chromatic effect for the SOPHIE instrument (Section 5.2) revealed intra-night FWHM and DRS RV variations up to 150 and 15 m s^{-1} respectively. This dramatic impact occurs when targets are observed over a wide airmass range, as is required in our high-cadence RV programme. I found that TERRA radial velocities were not significantly affected by this systematic effect, enabling SOPHIE RV precision of 1-2 m s^{-1} to be achieved in our programme. For HARPS observations, the only evidence for the chromatic effect was found in the FWHM values of an M-dwarf RV standard star; radial velocities were not affected throughout (Section 5.2.2). I found that RVs derived from TERRA showed significantly less scatter than the default values from DRS CCF computation, reaching stability below 1 m s^{-1} for our observations.

Planet discoveries from the RV programme

In Chapter 5 of this thesis I discussed the discovery of four new planet candidates around two F-type stars selected due to their anomalously low $\log(R'_{HK})$ values. They were observed as part of our RV programme with the HARPS

instrument. Observations of HD11231 revealed a hot Saturn-mass companion with an orbital period of approximately 6 days and stellar line profile distortions consistent with low amplitude γ-Doradus pulsations with unusually long periods (Section 5.3). Following WASP-118 (Močnik et al. 2017), this constitutes the second γ-Doradus planet host discovery published to date. A compact, low-mass multi-planet system orbits HD38677. The three planet candidates detected (HD38677 b, c and d) have periods of approximately 20, 10 and either 3 or 1.7 days and minimum masses of ~18, 13 and 3 $M_\oplus$ (Section 5.4). The alias ambiguity for HD38677 d is not yet resolved, and the 20 day signal should be validated with further data to test a possible stellar or longer-period origin.

Figure 6.1 shows the activity - colour distribution of my MS sample, using $\log(R'_{\rm HKmean})$, and known planet hosts, as already shown in Section 3.4.1. I now highlight HD11231 and HD38677, the two key targets with short period planet discoveries discussed in this thesis. In the context of its anomalously low $\log(R'_{\rm HK})$ value, HD11231 joins the group of HJ hosts first identified by Fossati et al. (2013), and extended by Staab et al. (2017) (c.f. Section 3.4.1). HD38677 joins the low-mass, short-period multi-planet systems Kepler-25 and Kepler-68 with below-basal $\log(R'_{\rm HK})$ that were first noted in Staab et al. (2017).

In the context of the known exoplanet population's mass-period distribution (Figure 1.1), HD11231 b is a fairly unremarkable Hot Jupiter type planet with a minimum mass between that of Saturn and Jupiter. HD11231 joins the ranks of the approximately 1% of FGK stars (Wang et al. 2015) hosting a short period gas giant. The three HD38677 planet candidates fall among the cluster of super-Earth and Neptune type planets on short to medium periods with measured $M_{\rm p}\ \sin i$ values. They are not particularly close the the sub-Jovian desert, which would have been circumstantial evidence for a recent evolution from inside the desert to its lower boundary via high mass loss rates.

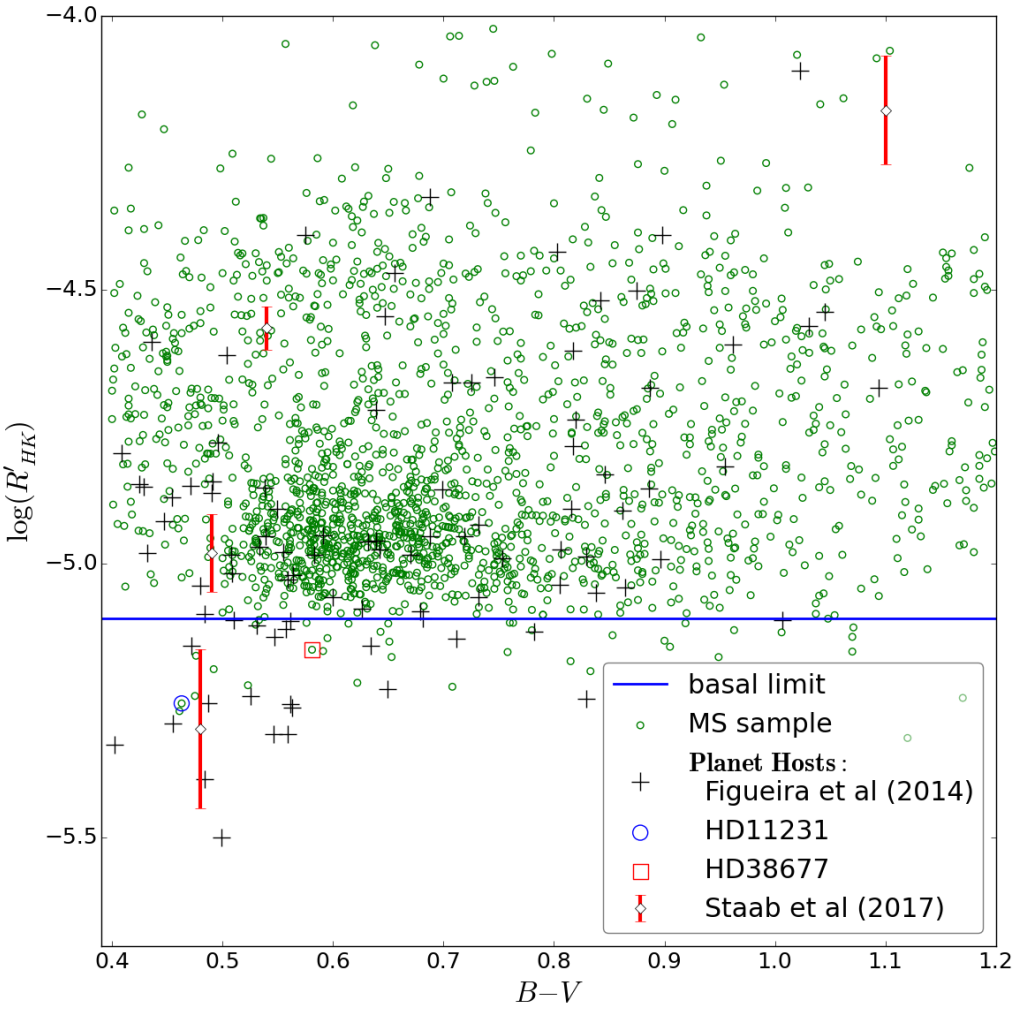

Figure 6.1: $\log(R'_{\mathrm{HKmean}})$ distribution for my MS sample (described in Section 2.3) compared to planet host data in the literature, HD11231 and HD38677.

Studies of the exoplanet occurrence rate (accounting for detection efficiency and selection bias) have shown that low mass planets with periods shorter than 50 days are relatively common (e.g. Howard et al. 2010). Approximately 40% of solar-type stars host at least one such planet with $M_{\mathrm{p}} \sin i < 30$ $\mathrm{M}_{\oplus}$ (Mayor et al. 2011). This HARPS-based study also showed that low-mass planets occur frequently in compact multi-planet systems, and that occurrence rates drop dramatically for periods below ~10 days. Due to the small number of known low-mass RV planets with P $\lesssim$ 5 days, there are large occurrence rate uncertainties for individual bins in this region of the $M_{\mathrm{p}} \sin i$ - period plane.

The plethora of transit detections from Kepler has allowed a more detailed insight into occurrence rates as a function of planet radius and orbital period. Fressin et al. (2013) confirmed a steep decline in the occurrence of super-Earth (1.25 - 2 $R_\oplus$) and sub-Neptune (2 - 4 $R_\oplus$) planets at periods below 10 days (c.f. Figure 10 in Sanchis-Ojeda et al. 2014) . Over the period range 0.8 - 3.4 days, only 0.83 ± 0.12% of stars host super-Earths (Fressin et al. 2013). In contrast, 5.3 ± 0.6% of stars have a sub-Neptune companion with periods between 17 and 29 days. The radii of the three HD38677 planet candidates are unknown, complicating direct comparisons with these occurrence rates. Their categorisation as a super-Earth (P = 3 or 1.7 days) and (sub-)Neptunes (P = 10 and 20 days) based on their $M_p \sin i$ values can be used to provide some approximate context for the discoveries. As discussed above, the presence of such low mass planets in a compact system is not particularly unusual in itself - see also Figure 7 in Winn & Fabrycky 2015 illustrating Kepler multi-planet systems. However, the innermost super-Earth does have a particularly short period when compared with planet occurrence rates. HD38677 is among the less than 1% of solar-type stars with such a short period low-mass companion, regardless of which alias solution is assumed.

As described in Section 5.1, our ongoing RV planet search has reached a phase where data is available for a significant number of targets but follow-up is still needed to reach firm conclusions for the vast majority of them. It is promising that we have already detected rare, short period planets for 2 out of only 17 targets with more than 22 RV observations (Section 5.1). For a full assessment of our RV program's detection rate follow-up observations and analysis must be completed, including an assessment of detection thresholds reached for each target in the $M_p \sin i$ - P plane. This will allow a rigorous statistical comparison of our detection efficiency with short period planet occurrence rates from the literature, and a final conclusion on the effectiveness of our strategy to target stars with anomalously low $\log(R'_{HK})$.

A preliminary assessment is possible based on the two planetary systems presented in this thesis. For this estimate, I conservatively assume that HD38677 and HD11231 remain the only targets (out of the sample of 17) for which short-period planets are discovered. Using published planet occurrence rates we can calculate the probability of detecting HD11231b and HD38677b, i.e. 1 Hot Jupiter and 1 short period super-Earth in a sample of 17 targets, under the assumption that the sample is unbiased towards the presence of short period planets.

The probability of a randomly selected star hosting a Hot Jupiter (P $\leq$ 10 d, $M_{\rm p} \sin i \geq 0.1$ M_J) is $p_{\rm HJ}$ = 1.2%, using the most recent precise RV occurrence rate published by Wright & Gaudi (2012). As already noted, the probability of detecting a super-Earth (1.25 - 2 $R_\oplus$) with a period below 3.4 days is $p_{\rm SE}$ = 0.83% for an unbiased sample. Fressin et al. (2013) also showed that over the sub-Neptune to Neptune regime (2 - 4 $R_\oplus$), for periods less than 3.4 days, $p_{\rm SE}$ = 0.21%. While the radius of HD38677b is unknown, we can put constraints on it based on the known exoplanet population. Rocky exoplanets with precise mass and radius data have been shown to fall on a tight $M_{\rm p} - R_{\rm p}$ relation consistent with Earth and Venus values (Dressing et al. 2015). Using this empirical relation, the expected minimum radius of HD38677b with $M_{\rm p} \sin i$ = 3 $M_\oplus$ is 1.4 $R_\oplus$. However we cannot rule out a low density and large radius for HD38677b, as found, e.g., in the Kepler-11 system (Lissauer et al. 2013). With a mass of 3 $M_\oplus$, planet radii of order 4 $R_\oplus$ are a possibility. I will therefore also take into account this scenario of a Neptune-size super-Earth planet.

The probability of a null result, defined as a target hosting neither a Hot Jupiter nor a short period super-Earth, is p_0 = 1-$p_{\rm HJ}$-$p_{\rm SE}$. I neglect the possible outcome of both types of planet being detected for the same target. This is an extremely rare planetary system architecture with only one case discovered to date (WASP-47, Huang et al. 2016) and a relatively poorly constrained occurrence of $1.0^{+9.5}_{-1.0}$% among HJ systems. Hot Jupiters are thought to reduce the

occurrence of other short period planets in a system, so the probabilities of both types of detection are not independent for a given target.

With these definitions, the probability of discovering 1 Hot Jupiter and 1 short period (P < 3.4 d) super-Earth with 1.25 - 2 $R_\oplus$ in an unbiased sample of 17 targets (and 15 null results) can be calculated[1] via the multinomial distribution as $17!/(1! \times 1! \times 15!)\ p_{HJ}\ p_{SE}\ p_0^{15}$ = 2.0%. If we instead assume that HD38677b has a radius in the sub-Neptune to Neptune regime (2 - 4 $R_\oplus$), then p_{SE}= 0.21%. The overall probability of detecting 1 such low density super-Earth and 1 Hot Jupiter is then 0.6%, so the former case is more conservative.

Such low probabilities do suggest that our sample of 17 targets is in fact biased towards the presence of short-period planets, presumably due to our selection based on anomalously low activity values. As the RV programme is continued, several preliminary candidate signals found for the remaining 15 stars could become short period planet detections. A stronger statement could already be made if we were to detect only one further short period super-Earth (p_{SE}= 0.83%) for another target, and establish 14 null results. The probability of such an outcome for an unbiased sample of 17 targets using multinomial statistics is only $17!/(1! \times 2! \times 14!)\ p_{HJ}\ p_{SE}^2\ p_0^{14}$ =0.13%.

[1] In these calculations we implicitly assume the sample has the same biases as the RV and Kepler samples from which the probabilites p_{HJ} and p_{SE} are derived.

6.2 Future work

Extending the RSS programme

In this thesis, I presented proof-of-concept stellar activity measurements using the RSS at SALT and its calibration to the Mount Wilson system. This work can now be exploited to efficiently measure $\log(R'_{\rm HK})$ values of faint Southern Hemisphere stars. Observations to achieve this for the vast majority of Southern transiting short period planet hosts are underway, and in total require less than 20 hours of SALT time. This programme will approximately double the number of known transiting HJ hosts with measured activity levels.

The $\log(R'_{\rm HK})$ values of the hosts of disintegrating rocky ultra short period planets would be of particular interest. Unfortunately only one (K2-22) out of three currently known hosts in this category is observable using SALT. K2-22 is an M-dwarf star, for which the calibration between S-values and $\log(R'_{\rm HK})$ is not well defined.

The RSS-derived database of $\log(R'_{\rm HK})$ values will be homogeneous, and not affected by the low-level systematic effects seen when comparing data from different instruments. Unlike most publications in the literature, the database will include a full uncertainty budget, as presented for the pilot programme in this thesis and Staab et al. (2017). In Section 3.3.3 I highlighted that current uncertainties are dominated by the calibration to the Mount Wilson system. A large number of additional calibrator star observations would reduce this uncertainty contribution for all science targets, including the four hosts characterised in this thesis.

Further systems with evidence for SPI or circumstellar absorption will likely be discovered in the upcoming extended RSS sample, covering a wide range of planetary and stellar properties. This will allow identification of systems for follow-up studies of SPI and planetary mass-loss. Combined with theoretical modelling, a better understanding of these processes and under which

circumstances they are important could be achieved. In the framework of characterising the correlation between $\log(R'_{\rm HK})$ and planetary surface gravity, the extended sample could provide constraints on the average effective stellar flux powering planetary mass loss (Fossati et al. 2015b).

A more comprehensive future investigation should aim to compare robust age estimates for all short period planet hosts to their activity levels. Stellar ages are notoriously difficult to determine, but such information would allow identification of further cases where the activity level is anomalous relative to the stellar age. The asteroseismic age determinations from PLATO 2.0 (Rauer et al. 2014 and references therein) should determine ages to 10% or better. With robust relative age values, one could more easily pick out systems where mass loss appears to be masking the intrinsic activity, but where the observed $\log(R'_{\rm HK})$ value is still above the basal limit. Systems where the stellar activity is boosted by SPI would also become more obvious.

HD38677 and HD11231

The uncertainties on the Keplerian parameters of the planetary systems discovered in this thesis should be rigorously quantified in a Bayesian framework. An obvious next step is to search for transits of the planets discovered in this work via the RV method. Using the systems' parameters from Tables 5.5, 5.6 and 5.9, the *a priori* transit probabilities are 13% for HD11231 b and 3%, 5% and 12% for HD38677 b, c and d. Since the stellar radius of HD38677 is unknown, I used a canonical 1.1 $R_\odot$ value for late F-type stars from Binney & Merrifield (1998) for these estimates. Note that if the alias solution at 1.7 days is assumed for HD38677 d, the transit probability increases to 18%. Either the detection of transits or further RV observations should resolve the alias ambiguity for HD38677 d. If any of the planets around HD11231 and HD38677 are found to be transiting, intensive follow-up studies via transmission spectroscopy should be done. Space-based near and far UV studies should provide the ultimate

confirmation of the hypothesis that these planets are undergoing significant mass loss, and would allow compositional analysis of the escaping material.

For HD11231, precision photometry over timescales of tens of days should further characterise the stellar pulsations, or at least provide upper limits for their photometric amplitude. High cadence precision photometry may be more effective than RV observations at identifying any additional, low-mass short period planets (interior to HD11231 b) for this system, depending on the amplitude ratio between photometric and RV pulsations.

Clearly, observations in the near and far UV with the Hubble space telescope are of high interest, even if the planets do not transit. These would reveal the true (unabsorbed) intrinsic activity level via UV emission lines, as done for WASP-13 in Fossati et al. (2015c). Time-resolved observations around inferior conjunction could reveal density increases in the circumstellar material, and constrain its composition. In the best case, the observations could reveal clear transit-like signals of the extended exospheres of the planets. The planets themselves would not necessarily have to transit the stars for this to occur: see the cases of 55 Cnc b in Ehrenreich et al. 2012 and GJ 436b (Ehrenreich et al. 2015), discussed in Section 1.1.2 of this thesis.

For HD11231 and HD38677, the presence of additional low-mass companions interior to the planets identified would be of interest. If present, these may contribute to the systems' circumstellar shrouding, or even provide the dominant mass loss source. Searches for additional planets in these systems could be carried out with both the transit and radial velocity methods.

Extending the study of anomalous activity values and the planet search programme

Clearly, an urgent priority is to complete the RV follow-up to a point where either planet detections or robust upper limits (null results) can be published. At that point it will be possible to carry out a full, rigorous assessment of the

short period planet detection rate resulting from the new approach of selecting targets with anomalously low $\log(R'_{\rm HK})$, and to compare this with the efficiency of traditional RV surveys. Going forward, our RV survey could be extended to all stars with below-basal activity levels in the MS sample presented in this thesis. All such objects are promising targets for space-based transit searches in the near future via the TESS (Ricker et al. 2014) and CHEOPS (Fortier et al. 2014) missions.

Additional targets with anomalously low $\log(R'_{\rm HK})$ could be identified in the P13 sample discussed in this thesis by examining outliers below the basal envelope of more evolved stars. The existence of such objects was suggested by Figure 3.9b, and a rigorous assment of the subgiant basal limit could follow the work in Mittag et al. (2013). Beyond the P13 sample, newly published activity catalogues should be searched for anomalously low values. In particular, two large databases have recently become public: Hempelmann et al. (2016) and Butler et al. (2017). The latter dataset is associated with an extensive release of HIRES radial velocity measurements, and is therefore particularly interesting for planet search follow-up.

Due to the format of the P13 compilation used in this thesis, I chose a conservative outlier selection criterion using the maximum activity value reported for each star. Additional outliers could be identified if the mean activity levels are used instead. In this case, a more robust approach than using the mean of $S_{\rm max}$ and $S_{\rm min}$ from P13 should be used due to the spurious $S_{\rm min}$ outliers I identified. A re-compilation of the catalogues underlying the P13 sample (Table 2.1) should be done, rejecting the spurious outliers from Isaacson & Fischer (2010), and using median S-values. Such a re-compilation effort could also account for small offsets between catalogues arising from the calibration to the Mount Wilson system, as identified by Jenkins et al. (2011). In addition, the apparent low-level dependence of the basal limit on stellar metallicity highlighted in the literature (Lovis et al. 2011) could be taken into account for the subset of

stars with accurate metallicity measurements. For such stars, the distinction between main sequence and evolved objects can also be made more accurately (e.g. following Jenkins et al. 2008). All of these steps should sharpen the basal activity envelope and reveal additional outliers more clearly.

As discussed in Section 2.2.1 of this thesis, the potential for exocomet transits to cause short-lived apparent activity anomalies should be investigated further. Detailed spectral analysis of these temporary activity drop-outs should be compared with systems like β Pictoris, and with any sungrazing comet transits found in the ongoing HARPS-N solar observations. Ultimately the prevalence and orbital dynamics of star-grazing comets could be investigated across several stellar systems. Extensive activity timeseries such as those in Butler et al. (2017) are needed for such studies.

Bibliography

ALMA Partnership et al., 2015, ApJ, 808, L1

Adams F. C., Laughlin G., 2006, ApJ, 649, 1004

Adams E. R., Jackson B., Endl M., 2016, AJ, 152, 47

Aigrain S., Hodgkin S., Irwin J., Hebb L., Irwin M., Favata F., Moraux E., Pont F., 2007, MNRAS, 375, 29

Aigrain S., Pont F., Zucker S., 2012, MNRAS, 419, 3147

Anderson E., Francis C., 2012, Astron. Lett., 38, 331

Anglada-Escudé G., Butler R. P., 2012, ApJS, 200, 15

Anglada-Escudé G., Tuomi M., 2012, A&A, 548, A58

Anglada-Escude G., Tuomi M., 2015, Science, 347, 1080

Anglada-Escudé G., et al., 2012, ApJ, 751, L16

Anglada-Escudé G., et al., 2013, A&A, 556, A126

Anglada-Escude G., et al., 2014, MNRAS Lett., 443, L89

Anglada-Escudé G., et al., 2016a, Nature, 536, 437

Anglada-Escudé G., et al., 2016b, ApJ, 830, 74

Arriagada P., 2011, ApJ, 734, 70

Arriagada P., Anglada-Escudé G., Butler R. P., Crane J. D., Shectman S. a., Thompson I., Wende S., Minniti D., 2013, ApJ, 771, 42

Bailey J. I., Mateo M. L., White R. J., Crane J. D., Shectman S. A., Team M. I., 2017, American Astronomical Society, AAS Meeting #229, id.403.03

Balachandran S., 1990, ApJ, 354, 310

Baliunas S. L., et al., 1995, ApJ, 438, 269

Balona L. A., 2015, MNRAS, 447, 2714

Baluev R. V., 2008, MNRAS, 385, 1279

Baluev R. V., 2009, MNRAS, 393, 969

Baranne A., et al., 1996, A&AS, 119, 373

Barge P., et al., 2008, A&A, 482, L17

Barnes J. R., Haswell C. A., Staab D., Anglada-Escudé G., 2016, MNRAS, 462, 1012

Batalha N. M., 2014, Proceedings of the National Academy of Sciences, 111, 12647

Beaugé C., Nesvorný D., 2013, ApJ, 763, 12

Beck P. G., et al., 2016, preprint (arXiv:1611.01402)

Ben-Jaffel L., Ballester G. E., 2013, A&A, 553, A52

Ben-Jaffel L., Ballester G. E., 2014, ApJ, 785, L30

Berdinas Z. M., Amado P. J., Anglada-Escudé G., Rodríguez-López C., Barnes J., 2016, MNRAS, 459, 3551

Berdinas Z. M., Rodríguez-López C., Amado P. J., Anglada-Escudé G., Barnes J. R., MacDonald J., Zechmeister M., Sarmiento L. F., 2017, accepted to MNRAS

Bernstein R., Shectman S. A., Gunnels S. M., Mochnacki S., Athey A. E., 2003, Proc. SPIE, 4841, 1694

Bessell M. S., Castelli F., Plez B., 1998, A&A, 333, 231

Beust H., Lagrange A.-M., Crawford I. A., Goudard C., Spyromilio J., Vidal-Madjar A., 1998, A&A, 338, 1015

Binney J., Merrifield M., 1998, Galactic Astronomy. Princeton University Press, Princeton, NJ

Birkby J. L., et al., 2014, MNRAS, 440, 1470

Bisikalo D., Kaygorodov P., Ionov D., Shematovich V., Lammer H., Fossati L., 2013, ApJ, 764, 19

Bochinski J. J., Haswell C. A., Marsh T. R., Dhillon V. S., Littlefair S. P., 2015, ApJ, 800, L21

Boesgaard A. M., Armengaud E., King J. R., 2003, ApJ, 582, 410

Boisse I., et al., 2009, A&A, 495, 959

Boisse I., et al., 2010, A&A, 523, A88

Boisse I., et al., 2013, A&A, 558, A86

Bonfanti A., Ortolani S., Piotto G., Nascimbeni V., 2015, A&A, 575, A18

Borgniet S., et al., 2014, A&A, 561, A65

Borgniet S., Meunier N., Lagrange A.-M., 2015, A&A, 581, A133

Borgniet S., Lagrange A.-M., Meunier N., Galland F., 2017, A&A, 599, A57

Bouchy F., et al., 2005, A&A, 444, L15

Bouchy F., Isambert J., Lovis C., Boisse I., Figueira P., Hébrard G., Pepe F., 2009a, EAS Publ. Series, 37, 247

Bouchy F., et al., 2009b, A&A, 496, 527

Bouchy F., et al., 2009c, A&A, 505, 853

Bouchy F., Diaz R. F., Hébrard G., Arnold L., Boisse I., Delfosse X., Perruchot S., Santerne A., 2012, A&A, 49, 12

Bourrier V., Hébrard G., 2014, A&A, 569, A65

Bourrier V., Ehrenreich D., Lecavelier des Etangs A., 2015, A&A, 582, A65

Bradley P. A., Guzik J. A., Miles L. F., Uytterhoeven K., Jackiewicz J., Kinemuchi K., 2015, AJ, 149, 68

Breger M., 1979, PASP, 91, 5

Brown T. M., Gilliland R. L., Noyes R. W., Ramsey L. W., 1991, ApJ, 368, 599

Brown D. J. a., et al., 2012, MNRAS, 423, 1503

Buccino A. P., Mauas P. J. D., 2008, A&A, 483, 903

Buchholz B., Ulmschneider P., Cuntz M., 1998, ApJ, 494, 700

Buckley D. A. H., Swart G. P., Meiring J. G., 2006, Proc. SPIE, 6267

Budaj J., 2013, A&A, 557, A72

Burrows A. S., 2014, Proceedings of the National Academy of Sciences, 111, 12601

Burton J. R., Watson C. A., Fitzsimmons A., Pollacco D., Moulds V., Littlefair S. P., Wheatley P. J., 2014, ApJ, 789, 113

Busuttil R., 2017, PhD thesis, The Open University

Butler R. P., Marcy G. W., Williams E., McCarthy C., Dosanjh P., Vogt S. S., 1996, PASP, 108, 500

Butler R. P., et al., 2017, AJ, 153, 208

Cameron A. C., et al., 2010, MNRAS, 407, 507

Canto Martins B. L., das Chagas M. L., Alves S., Le~ao I. C., de Souza Neto L. P., de Medeiros J. R., 2011, A&A, 530, A73

Carilli C., Rawlings S., 2004, New Astronomy Reviews, 48, 979

Casagrande L., Schönrich R., Asplund M., Cassisi S., Ramírez I., Meléndez J., Bensby T., Feltzing S., 2011, A&A, 530, A138

Cauley P. W., Redfield S., Jensen A. G., Barman T., Endl M., Cochran W. D., 2015, ApJ, 810, 13

Cauley P. W., Redfield S., Jensen A. G., 2017, preprint (arXiv:1703.09543)

Chadney J., Galand M., Unruh Y., Koskinen T., Sanz-Forcada J., 2015, Icarus, 250, 357

Charbonneau D., Brown T. M. T., Noyes R. R. W. R., Gilliland R. L. R., 2002, ApJ, 568, 377

Chen G., et al., 2014, A&A, 563, A40

Cincunegui C., Díaz R. F., Mauas P. J. D., 2007, A&A, 469, 309

Cochran W. D., Hatzes A. P., Paulson D. B., 2002, AJ, 124, 565

Cohen O., Kashyap V. L., Drake J. J., Sokolov I. V., Garraffo C., Gombosi T. I., 2011, ApJ, 733, 67

Cosentino R., et al., 2012, Proc. SPIE, 8446

Cosentino R., et al., 2014, Proc. SPIE, 9147

Crawford S. M., et al., 2010, Proc. SPIE, 7737

Crockett C. J., Mahmud N. I., Prato L., Johns-Krull C. M., Jaffe D. T., Hartigan P. M., Beichman C. a., 2012, ApJ, 761, 164

Cumming A., 2004, MNRAS, 354, 1165

Cunha D., Figueira P., Santos N. C., Lovis C., Boué G., 2013, A&A, 550, A75

Cuntz M., Saar S. H., Musielak Z. E., 2000, ApJ, 533, L151

Cutri R. M., Al. E., 2012, VizieR On-line Data Catalog: II/311, 2311

Cutri R. M., et al., 2003, VizieR On-line Data Catalog: II/246., 2246

Czesla S., Salz M., Schneider P. C., Schmitt J. H. M. M., 2013, A&A, 560, A17

David T. J., et al., 2016, AJ, 151, 112

Davie M. W., Pollard K. R., Cottrell P. L., Brunsden E., Wright D. J., De Cat P., 2014, Publications of the Astronomical Society of Australia, 31, e025

Dawson R. I., Fabrycky D. C., 2010, ApJ, 722, 937

De Silva G. M., Sneden C., Paulson D. B., Asplund M., Bland-Hawthorn J., Bessell M. S., Freeman K. C., 2006, AJ, 131, 455

Demory B.-O., Gillon M., Madhusudhan N., Queloz D., 2016, MNRAS, 455, 2018

Desort M., Lagrange A.-M., Galland F., Udry S., Mayor M., 2007, A&A, 473, 983

Díaz R. F., et al., 2012, A&A, 538, A113

Dolphin A. E., 2000, PASP, 112, 1397

Douglas S. T., et al., 2014, ApJ, 795

Doyle L. R., et al., 2011, Science, 333, 1602

Dressing C. D., et al., 2015, The Astrophysical Journal, 800, 135

Dumusque X., Udry S., Lovis C., Santos N. C., Monteiro M. J. P. F. G., 2011a, A&A, 525, A140

Dumusque X., Santos N. C., Udry S., Lovis C., Bonfils X., 2011b, A&A, 527, A82

Dumusque X., et al., 2012, Nature, 491, 207

Dumusque X., et al., 2014, ApJ, 789, 154

Dumusque X., et al., 2015, ApJ, 814, L21

Duncan D. K., et al., 1991, ApJS, 76, 383

Dupret M., 2007, Communications in Asteroseismology, 150, 98

Ehrenreich D., et al., 2012, A&A, 18, 4

Ehrenreich D., et al., 2015, Nature, 522, 459

Enoch B., et al., 2011, AJ, 142, 86

Erkaev N. V., Kulikov Y. N., Lammer H., Selsis F., Langmayr D., Jaritz G. F., Biernat H. K., 2007, A&A, 472, 329

Ferlet R., Hobbs L. M., Vidal-Madjar A., 1987, A&A, 185, 267

Figueira P., Santos N. C., Pepe F., Lovis C., Nardetto N., 2013, A&A, 557, A93

Figueira P., Oshagh M., Adibekyan V. Z., Santos N. C., 2014, A&A, 572, A51

Figueira P., et al., 2016, A&A, 592, A143

Fischer D. A., et al., 2008, ApJ, 675, 790

Fischer D. A., et al., 2016, PASP, 128, 066001

Fortier A., Beck T., Benz W., Broeg C., Cessa V., Ehrenreich D., Thomas N., 2014, Proc. SPIE, 9143

Fossati L., et al., 2010a, ApJ, 714, L222

Fossati L., et al., 2010b, ApJ, 720, 872

Fossati L., Ryabchikova T., Shulyak D. V., Haswell C. A., Elmasli A., Pandey C. P., Barnes T. G., Zwintz K., 2011, MNRAS, 417, 495

Fossati L., Ayres T. R., Haswell C. a., Bohlender D., Kochukhov O., Flöer L., 2013, ApJ, 766, L20

Fossati L., et al., 2015a, preprint (arXiv:1503.01278)

Fossati L., Ingrassia S., Lanza a. F., 2015b, ApJ, 812, L35

Fossati L., France K., Koskinen T., Juvan I. G., Haswell C. A., Lendl M., 2015c, ApJ, 815, 118

Fossati L., et al., 2017, accepted to A&A (forthcoming article)

Fressin F., et al., 2013, ApJ, 766, 81

Gaia Collaboration et al., 2016, A&A, 595, A1

Gaidos E., et al., 2017, MNRAS, 464, 850

Galland F., Lagrange A.-M., Udry S., Chelli A., Pepe F., Queloz D., Beuzit J.-L., Mayor M., 2005, A&A, 443, 337

Galland F., Lagrange A.-M., Udry S., Beuzit J.-L., Pepe F., Mayor M., 2006, A&A, 452, 709

Gaudi B. S., 2012, Annual Review of Astronomy and Astrophysics, 50, 411

Gilliland R. L., et al., 2013, ApJ, 766, 40

Gillon M., et al., 2013, A&A, 552, A82

Gillon M., et al., 2014, A&A, 562, L3

Gillon M., et al., 2017, Nature, 542, 456

Gomes da Silva J., Santos N. C., Boisse I., Dumusque X., Lovis C., 2014, A&A, 566, A66

Gorti U., Liseau R., Sándor Z., Clarke C., 2016, Space Sci. Rev., 205, 125

Grassitelli L., Fossati L., Langer N., Miglio A., Istrate A. G., Sanyal D., 2015, A&A, 584, L2

Gray R. O., Corbally C. J., Garrison R. F., McFadden M. T., Robinson P. E., 2003, AJ, 126, 2048

Gray R. O., Corbally C. J., Garrison R. F., McFadden M. T., Bubar E. J., McGahee C. E., O'Donoghue a. a., Knox E. R., 2006, AJ, 132, 161

Guenther E. W., et al., 2011, A&A, 525, A24

Halbwachs J. L., Mayor M., Udry S., Arenou F., 2003, A&A, 397, 159

Hall J. C., 2005, Living Reviews in Solar Physics, 5

Hall J. C., Fulton E. E., Huenemoerder D. P., Welty A. D., Neff J. E., 1994, PASP, 106, 315

Hall J. C., Lockwood G. W., Skiff B. a., 2007, AJ, 133, 862

Han E., Wang S. X., Wright J. T., Feng Y. K., Zhao M., Fakhouri O., Brown J. I., Hancock C., 2014, PASP, 126, 827

Hartman J. D., 2010, ApJ, 717, L138

Hartman J. D., Bakos G. Á., Kovács G., Noyes R. W., 2010, MNRAS, 408, 17

Haswell C. a., et al., 2012, ApJ, 760, 79

Hatzes a. P., 2014, A&A, 568, A84

Hay K. L., et al., 2016, MNRAS, 463, 3276

Haywood R. D., et al., 2014, MNRAS, 443, 2517

Hebb L., et al., 2009, ApJ, 693, 1920

Hellier C., et al., 2011, A&A, 535, L7

Hempelmann A., Mittag M., Gonzalez-Perez J. N., Schmitt J. H. M. M., Schröder K. P., Rauw G., 2016, A&A, 586, A14

Henry G. W., Fekel F. C., 2005, AJ, 129, 2026

Henry T. J., Soderblom D. R., Donahue R. a., Baliunas S. L., 1996, AJ, 111, 439

Henry G. W., Fekel F. C., Henry S. M., 2011, AJ, 142, 39

Hernán-Obispo M., Gálvez-Ortiz M. C., Anglada-Escudé G., Kane S. R., Barnes J. R., de Castro E., Cornide M., 2010, A&A, 512, A45

Hernán-Obispo M., et al., 2015, A&A, 576, A66

Herrero E., Morales J. C., Ribas I., Naves R., 2011, A&A, 526, L10

Hillenbrand L., Isaacson H., Marcy G., Barenfeld S., Fischer D., Howard A., 2014, preprint, 18, 759 (arXiv:1408.3475)

Hoang T., Lazarian A., Burkhart B., Loeb A., 2017, ApJ, 837, 5

Hobbs L. M., Vidal-Madjar A., Ferlet R., Albert C. E., Gry C., 1985, ApJ, 293, 29

Horne K., 1986, PASP, 98, 609

Howard A. W., et al., 2010, Science, 330, 653

Howell S. B., et al., 2014, PASP, 126, 398

Huang C., Wu Y., Triaud A. H. M. J., 2016, The Astrophysical Journal, 825, 98

Isaacson H., Fischer D., 2010, ApJ, 725, 875

Iyer A. R., Swain M. R., Zellem R. T., Line M. R., Roudier G., Rocha G., Livingston J. H., 2016, ApJ, 823, 109

Jackson B., Greenberg R., Barnes R., 2008, ApJ, 678, 1396

Jenkins J. S., et al., 2006, MNRAS, 372, 163

Jenkins J. S., Jones H. R. a., Pavlenko Y., Pinfield D. J., Barnes J. R., Lyubchik Y., 2008, A&A, 485, 571

Jenkins J. S., et al., 2011, A&A, 531, A8

Jin S., Mordasini C., Parmentier V., van Boekel R., Henning T., Ji J., 2014, ApJ, 795, 65

Johnson J. A., et al., 2011, ApJ, 735, 24

Jones H. R. A., Butler R. P., Tinney C. G., Marcy G. W., Carter B. D., Penny A. J., McCarthy C., Bailey J., 2006, MNRAS, 369, 249

Kataria T., Showman A. P., Fortney J. J., Stevenson K. B., Line M. R., Kreidberg L., Bean J. L., Désert J.-M., 2015, ApJ, 801, 86

Kaye A. B., Handler G., Krisciunas K., Poretti E., Zerbi F. M., 1999, PASP, 111, 840

Kharchenko N. V., Piskunov A. E., Röser S., Schilbach E., Scholz R.-D., Zinnecker H., 2009, A&A, 504, 681

Kiefer F., des Etangs A. L., Boissier J., Vidal-Madjar A., Beust H., Lagrange A.-M., Hébrard G., Ferlet R., 2014a, Nature, 514, 462

Kiefer F., Lecavelier des Etangs A., -C Augereau J., Vidal-Madjar A., Lagrange A.-M., Beust H., 2014b, A&A, 561, L10

Kislyakova K. G., et al., 2016, MNRAS, 461, 988

Kite E. S., Bruce Fegley Jr. Schaefer L., Gaidos E., 2016, ApJ, 828, 80

Knight M. M., 2008, PhD thesis, University of Maryland, http://adsabs.harvard.edu/abs/2008PhDT.....14K

Knutson H. a., Howard A. W., Isaacson H., 2010, ApJ, 720, 1569

Kobulnicky H. A., Nordsieck K. H., Burgh E. B., Smith M. P., Percival J. W., Williams T. B., O'Donoghue D., 2003, Proc. SPIE, 4841

Konacki M., Muterspaugh M. W., Kulkarni S. R., Hełminiak K. G., 2009, ApJ, 704, 513

Koskinen T., Harris M., Yelle R., Lavvas P., 2013, Icarus, 226, 1678

Kreidberg L., et al., 2014, ApJ, 793, L27

Krejčová T., Budaj J., 2012, A&A, 540, A82

Kroupa P., 1998, MNRAS, 298, 231

Kurokawa H., Nakamoto T., 2014, ApJ, 783, 54

Kurucz R. L., Furenlid I., Brault J., Testerman L., 1984, National Solar Observatory Atlas

Lagrange A.-M., Desort M., Galland F., Udry S., Mayor M., 2009, A&A, 495, 335

Lagrange A.-M., Meunier N., Chauvin G., Sterzik M., Galland F., Lo Curto G., Rameau J., Sosnowska D., 2013, A&A, 559, A83

Lammer H., Selsis F., Ribas I., Guinan E. F., Bauer S. J., Weiss W. W., 2003, ApJ, 598, L121

Lannier J., et al., 2016, A&A, 596, A83

Lanza A. F., 2008, A&A, 487, 1163

Lanza A. F., 2012, A&A, 544, A23

Lanza A. F., 2013, A&A, 557, A31

Lanza A. F., 2014, A&A, 572, L6

Lecavelier des Etangs A., et al., 2010, A&A, 514, A72

Lecavelier des Etangs a., et al., 2012, A&A, 543, L4

Lenz L. F., Reiners A., Kürster M., 2011, in 16th Cambridge Workshop on Cool Stars. http://adsabs.harvard.edu/abs/2011ASPC..448.1173L

Linsky J. L., Yang H., France K., Froning C. S., Green J. C., Stocke J. T., Osterman S. N., 2010, ApJ, 717, 1291

Lissauer J. J., et al., 2011, Nature, 470, 53

Lissauer J. J., et al., 2013, The Astrophysical Journal, 770, 131

Lissauer J. J., et al., 2014, ApJ, 784, 44

Lo Curto G., et al., 2015, The Messenger, vol. 162, p. 9-15, 162

Lockwood G. W., Skiff B. a., Henry G. W., Henry S., Radick R. R., Baliunas S. L., Donahue R. a., Soon W., 2007, ApJS, 171, 260

Lomb N. R., 1976, Physics, 39

Lopez E. D., Rice K., 2016, preprint (arXiv:1610.09390)

López-Morales M., et al., 2008, AJ, 136, 1901

López-Santiago J., Montes D., Gálvez-Ortiz M. C., Crespo-Chacón I., Martínez-Arnáiz R. M., Fernández-Figueroa M. J., de Castro E., Cornide M., 2010, A&A, 514, A97

Lopez E. D., Fortney J. J., Miller N., 2012, ApJ, 761, 59

Lovis C., et al., 2006, Nature, 441, 305

Lovis C., Fischer D. A., Dotson R., 2010, Radial Velocity Techniques for Exoplanets. University of Arizona Press, pp 27–54

Lovis C., et al., 2011, preprint (arXiv:1107.5325)

Lundkvist M. S., et al., 2016, Nature Communications, 7, 11201

Madhusudhan N., Lee K. K. M., Mousis O., 2012, ApJ, 759, L40

Maggio a., et al., 2015, ApJ, 811, L2

Maisonneuve F., et al., 2011, MNRAS, 415, 2977

Mamajek E. E., Hillenbrand L. a., 2008, ApJ, 687, 1264

Mann A. W., et al., 2016, ApJ, 818, 46

Mann A. W., et al., 2017, AJ, 153, 64

Mathias P., et al., 2004, A&A, 417, 189

Matsakos T., Königl A., 2016, ApJ, 820, L8

Matsakos T., Uribe A., Königl A., 2015, A&A, 578, A6

Mayor M., Queloz D., 1995, Nature, 378, 355

Mayor M., et al., 2003, The Messenger, No.114, p. 20-24, 114, 20

Mayor M., et al., 2009, A&A, 493, 639

Mayor M., et al., 2011, preprint (arXiv:1109.2497)

Mayor M., Lovis C., Santos N. C., 2014, Nature, 513, 328

Mermilliod J.-C., Rosvick J., 1992, A&A, 265, 513

Meschiari S., Wolf A. S., Rivera E., Laughlin G., Vogt S., Butler P., 2009, PASP, 121, 1016

Meunier N., Lagrange A.-M., Borgniet S., Rieutord M., 2015, A&A, 583, A118

Miguel Y., Kaltenegger L., Fegley B., Schaefer L., 2011, ApJ, 742, L19

Miller B. P., Gallo E., Wright J. T., Dupree A. K., 2012, ApJ, 754, 137

Miller B. P., Gallo E., Wright J. T., Pearson E. G., 2015, ApJ, 799, 163

Minniti D., Butler R. P., López-Morales M., Shectman S. a., Adams F. C., Arriagada P., Boss A. P., Chambers J. E., 2009, ApJ, 693, 1424

Mittag M., Schmitt J. H. M. M., Schröder K.-P., 2013, A&A, 549, A117

Mochejska B. J., Stanek K. Z., Sasselov D. D., Szentgyorgyi A. H., 2002, AJ, 123, 3460

Močnik T., Hellier C., Anderson D. R., Clark B. J. M., 2017, preprint (arXiv:1702.05078)

Montgomery S. L., Welsh B. Y., 2012, PASP, 124, 1042

Motalebi F., et al., 2015, A&A, 584, A72

Mura a., et al., 2011, Icarus, 211, 1

Murgas F., Pallé E., Zapatero Osorio M. R., Nortmann L., Hoyer S., Cabrera-Lavers A., 2014, A&A, 563, A41

Narain U., Ulmschneider P., 1996, Space Sci. Rev., 75, 453

Niedzielski A., et al., 2015, A&A, 573, A36

Noyes R. W., Hartmann L. W., Baliunas S. L., Duncan D. K., Vaughan A. H., 1984, ApJ, 279, 763

Pace G., 2013, A&A, 551, L8

Paulson D. B., Saar S. H., Cochran W. D., Hatzes A. P., 2002, AJ, 124, 572

Paulson D. B., Sneden C., Cochran W. D., 2003, AJ, 125, 3185

Paulson D. B., Cochran W. D., Hatzes A. P., 2004, AJ, 127, 3579

Pecaut M. J., Mamajek E. E., 2013, ApJS, 208, 9

Penny M. T., Rattenbury N. J., Gaudi B. S., Kerins E., 2017, AJ, 153, 161

Pepe F., et al., 2002a, The Messenger, No. 110, p. 9 - 14, 110, 9

Pepe F., Mayor M., Galland F., Naef D., Queloz D., Santos N. C., Udry S., Burnet M., 2002b, A&A, 388, 632

Pepe F., et al., 2003, Proc. SPIE, 4841, 1045

Pepe F., et al., 2011, A&A, 534, A58

Pepe F., et al., 2013, Nature, 503, 377

Perez-Becker D., Chiang E., 2013, MNRAS, 433, 2294

Perez Martinez M. I., Schroder K.-P., Hauschildt P., 2014, MNRAS, 445, 270

Perruchot S., et al., 2008, Proc. SPIE, 7014

Perryman M. A. C., et al., 1998, A&A, 120, 81

Perryman M., Hartman J., Bakos G. Á., Lindegren L., 2014, ApJ, 797, 14

Pillitteri I., Günther H. M., Wolk S. J., Kashyap V. L., Cohen O., 2011, ApJ, 741, L18

Pillitteri I., Wolk S. J., Sciortino S., Antoci V., 2014, A&A, 567, A128

Pollacco D. L., et al., 2006, PASP, 118, 1407

Pollacco D., et al., 2008, MNRAS, 385, 1576

Poppenhaeger K., Schmitt J. H. M. M., 2011, ApJ, 735, 59

Poppenhaeger K., Wolk S. J., 2014, A&A, 565, L1

Poppenhaeger K., Robrade J., Schmitt J. H. M. M., 2010, A&A, 515, A98

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of scientific computing. Cambridge: University Press, 2nd ed.

Queloz D., et al., 2001, A&A, 379, 279

Queloz D., et al., 2009, A&A, 506, 303

Quinn S. N., et al., 2014, ApJ, 787, 27

Rajpaul V., Aigrain S., Roberts S., 2016, MNRAS Lett., 456, L6

Rappaport S., et al., 2012, ApJ, 752, 1

Rappaport S., Sanchis-Ojeda R., Rogers L. a., Levine A., Winn J. N., 2013, ApJ, 773, L15

Rappaport S., Barclay T., DeVore J., Rowe J., Sanchis-Ojeda R., Still M., 2014, ApJ, 784, 40

Rauer H., et al., 2014, Exp. Astron., 38, 249

Redfield S., Linsky J. L., 2008, ApJ, 673, 283

Ricker G. R., et al., 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Ridden-Harper A. R., et al., 2016, A&A, 593, A129

Ritchey a. M., Martinez M., Pan K., Federman S. R., Lambert D. L., 2006, ApJ, 649, 788

Robertson P., Mahadevan S., 2014, ApJ, 793, L24

Robertson P., Mahadevan S., Endl M., Roy a., 2014, Science, 345, 440

Robertson P., Roy A., Mahadevan S., 2015, ApJ, 805, L22

Ryabchikova T., Fossati L., Shulyak D., 2009, A&A, 506, 203

Salpeter E. E., 1955, ApJ, 121, 161

Salz M., Czesla S., Schneider P. C., Schmitt J. H. M. M., 2016, A&A, 586, A75

Sanchis-Ojeda R., Rappaport S., Winn J. N., Kotson M. C., Levine A., Mellah I. E., 2014, ApJ, 787, 47

Sanchis-Ojeda R., et al., 2015, ApJ, 812, 112

Santerne a., et al., 2012, A&A, 545, A76

Santos N. C., et al., 2014, A&A, 566, A35

Sato B., et al., 2012, PASJ, 64, 135

Saur J., Grambusch T., Duling S., Neubauer F. M., Simon S., 2013, A&A, 552, A119

Scargle J. D., 1982, ApJ, 263, 835

Scharf C. A., 2010, ApJ, 722, 1547

Schilbach E., Robichon N., Souchay J., Guibert J., 1995, A&A, 299, 696

Schrijver C., 1987, A&A, 172, 111

Schrijver C. J., Zwaan C., 2000, Solar and stellar magnetic activity. New York : Cambridge University Press

Schrijver C. J., Dobson A. K., Radick R. R., 1989, ApJ, 341, 1035

Schröder C., Reiners a., Schmitt J. H. M. M., 2009, A&A, 1107, 1099

Schröder K.-P., Mittag M., Pérez Martínez M. I., Cuntz M., Schmitt J. H. M. M., 2012, A&A, 540, A130

Seabold S., Perktold J., 2010, in 9th Python in Science Conference.

See V., Jardine M., Fares R., Donati J.-F., Moutou C., 2015, MNRAS, 450, 4323

Sekiguchi M., Fukugita M., 2000, AJ, 120, 1072

Shetrone M. D., Rivera F., Smith G. H., Carveth C., 2008, PASP, 120, 730

Shkolnik E., Walker G. A. H., Bohlender D. A., 2003, ApJ, 597, 1092

Shkolnik E., Bohlender D. A., Walker G. A. H., Collier Cameron A., 2008, ApJ, 676, 628

Skumanich A., 1972, ApJ, 171, 565

Soderblom D. R., Nelan E., Benedict G. F., McArthur B., Ramirez I., Spiesman W., Jones B. F., 2005, AJ, 129, 1616

Soubiran C., Crifo F., Chemin L., 2015, Radial Velocity Standards : update, www.obs.u-bordeaux1.fr/m2a/soubiran/GAIA-C6-TN-LAB-CS-015.pdf

Southworth J., 2011, MNRAS, 417, 2166

Southworth J., et al., 2014, MNRAS, 447, 711

Staab D., Haswell C. A., Smith G. D., Fossati L., Barnes J. R., Busuttil R., Jenkins J. S., 2017, MNRAS, 466, 738

Stauffer J. R., et al., 2007, ApJS, 172, 663

Stephens M. a., 1974, Journal of the American Statistical Association, 69, 730

Stevenson K. B., et al., 2014, Science, 346, 838

Strassmeier K. G., Washuettl a., Granzer T., Scheck M., Weber M., 2000, A&AS, 142, 275

Strugarek A., Brun A. S., Matt S. P., Réville V., 2015, ApJ, 815, 111

Sumi T., et al., 2011, Nature, 473, 349

Szabó G. M., Kiss L. L., 2011, ApJ, 727, L44

Tabernero H. M., Montes D., González Hernández J. I., 2012, A&A, 547, A13

Takeda Y., Tajitsu A., Honda S., Kawanomoto S., Ando H., Sakurai T., 2012, PASJ, 64, 12

Takeda Y., Honda S., Ohnishi T., Ohkubo M., Hirata R., Sadakane K., 2013, PASJ, 65, 53

Taylor M. B., 2005, Astronomical Data Analysis Software and Systems XIV - ASP Conference Series, 347, 29

Tinney C. G., McCarthy C., Jones H. R. A., Butler R. P., Carter B. D., Marcy G. W., Penny A. J., 2002, MNRAS, 332, 759

Triaud a. H. M. J., et al., 2010, A&A, 524, 25

Tripathi A., et al., 2010, ApJ, 715, 421

Tuomi M., 2012, A&A, 543, A52

Tuomi M., Anglada-Escudé G., Gerlach E., Jones H. R. a., Reiners A., Rivera E. J., Vogt S. S., Butler R. P., 2012, A&A, 549, A48

Tuomi M., et al., 2013, A&A, 551, A79

Tuomi M., Jones H. R. A., Barnes J. R., Anglada-Escude G., Jenkins J. S., 2014, MNRAS, 441, 1545

Van Der Bliek N., Manfroid J., Bouchet P., 1996, A&AS, 119, 547

Vaughan A. H., Preston G. W., 1980, PASP, 92, 385

Vidal-Madjar A., Des Etangs A. L., Désert J.-M., Ballester G. E., Ferlet R., Hébrard G., Mayor M., 2003, Nature, 422, 143

Vidal-Madjar a., et al., 2013, A&A, 54, 1

Vidotto A. A., Fares R., Jardine M., Moutou C., Donati J.-F., 2015, MNRAS, 449, 4117

Wakeford H. R., et al., 2017, ApJ, 835, L12

Walker G. A. H., et al., 2008, A&A, 482, 691

Wang J., Fischer D. A., Horch E. P., Huang X., 2015, ApJ, 799, 229

Welsh B. Y., Montgomery S., 2013, PASP, 125, 759

Welsh B. Y., Lallement R., Vergely J.-L., Raimond S., 2010, A&A, 510, A54

Welsh W. F., et al., 2011, ApJS, 197, 4

White R. J., Gabor J. M., Hillenbrand L. a., 2007, AJ, 133, 2524

Wilson O. C., 1968, ApJ, 153, 221

Winn J. N., Fabrycky D. C., 2015, Annual Review of Astronomy and Astrophysics, 53, 409

Winn J. N., et al., 2011, ApJ, 737, L18

Wolszczan A., Frail D. A., 1992, Nature, 355, 145

Wong I., et al., 2015, ApJ, 811, 122

Wright J. T., 2004, AJ, 128, 1273

Wright J. T., 2005a, PASP, 117, 657

Wright J. T., 2005b, AJ, 129, 1776

Wright J. T., Eastman J. D., 2014, PASP, 126, 838

Wright J. T., Gaudi B. S., 2012, Exoplanet Detection Methods. p. 489 (arXiv:1210.2471)

Wright J. T., Howard A. W., 2009, ApJS, 182, 205

Wright J. T., Marcy G. W., Butler R. P., Vogt S. S., 2004, ApJ, 152, 261

Wright E. L., et al., 2010, AJ, 140, 1868

Wyatt M. C., Panić O., Kennedy G. M., Matrà L., 2015, Astrophys. Space Sci., 357, 103

Wyttenbach A., Ehrenreich D., Lovis C., Udry S., Pepe F., 2015, A&A, 577, A62

Zechmeister M., et al., 2013, A&A, 552, A78

da Silva L., et al., 2006, A&A, 458, 609

de Bruijne J. H. J., Hoogerwerf R., de Zeeuw P. T., 2001, A&A, 367, 111

de Wit J., et al., 2017, ApJ, 836, L17

van Leeuwen F., 2007, A&A, 474, 653

van Lieshout R., Min M., Dominik C., 2014, A&A, 572, A76

van Saders J. L., Gaudi B. S., 2011, ApJ, 729, 63

van de Kamp P., 1977, Vistas in Astronomy, 21, 289